Indicators of State Failure

Phase II

David Carment, Carleton University, Yiagadeesen Samy, Carleton University, Doug Hales, CAE Professional Services Canada, Jordan Miller, CAE Professional Services Canada, Liz St.Jean, CAE Professional Services Canada, Peter Tikuisis, DRDC.

CAE Professional Services
300-1135 Innovation Dr
Ottawa, ON
K2K 3G7

Defence R&D Canada – Toronto: Adversarial Intent Contract Report

DRDC Toronto CR 2010-124
W7711-088139-001-TOR
August 2010

Principal Author

Original signed by David Carment

David Carment

Carleton University

Approved by

Original signed by Peter Tikuisis

Peter Tikuisis

Contract Scientific Authority

Approved for release by

Original signed by K. C. Wulterkens

K. C. Wulterkens

For Chair, Knowledge and Information Management Committee

Defence R&D Canada – Toronto

Abstract

As noted in the Canada First Defence Strategy and reiterated in the more recent US Quarterly Defence Review, instability and state failure in distant lands can directly affect our own security and that of our allies. Development of a predictive model has become both a topical issue and an increasingly important area of research in academic and policy communities. This is the second report documenting CAE's support to DRDC's continuing efforts to develop an Early Warning Model (EWM) of state instability.

The conceptual framework for an EWM was developed in a previous project, though without a data set to validate assumptions and the general hypothesis. The focus of the current project was to collect and code events data and integrate it with structural data that will ultimately be used to calibrate and validate the conceptual model. A descriptive framework was established and incidences of failure identified using the methodology developed by Carleton University's Country Indicators for Foreign Policy (CIFP) project. Events data were then collected for the 24 months preceding these instances of state failure. Periods of relative stability for these states were also identified and events data collected for these periods. These events were distinguished as eroding or bolstering state Authority, Legitimacy or Capacity, including the severity of the challenge recorded. An inter-coder reliability test was conducted to confirm coding consistency. The results were compared with data available through Virtual Research Associates (VRA), thus affording an opportunity to gauge the merits of human (versus machine) coding. VRA provides a web-based software tool that supports interactive analysis and intuitive display of newswire data in the form of tables, graphs and charts. The visual displays represent threat warning for potential international 'hotspots'. The research effort concluded that human coding is more discriminating but also considerably more time consuming. An extensive data base has been developed and analysis commenced, which will continue beyond the submission of this report.

The conceptual model envisages integrating events and structural data which would allow for the measurement and monitoring of state tension and, through regression analysis, for vulnerability and instability thresholds to be determined and crises of interest to be forecast. This report documents a uniquely extensive data base that has been developed to support this effort.

Résumé

Comme il est indiqué dans la Stratégie de défense *Le Canada d'abord,* et plus récemment encore dans le *Quaterly Defence Review* aux États-Unis, l'instabilité et la mise en déroute d'un État à l'autre bout du monde peuvent nuire directement à notre propre sécurité et à celle de nos alliés. Le développement d'un modèle prédictif est devenu un problème d'actualité et un sujet de recherche de plus en plus important pour le milieu de l'enseignement et pour le secteur des politiques. Le présent rapport est le deuxième à documenter le soutien de l'IAO aux efforts constants de RDDC pour le développement d'un modèle de pré-alerte (EWM) de l'instabilité des États.

Le cadre conceptuel d'un EWM a été développé au cours d'un projet antérieur, malgré l'absence d'un ensemble de données pour valider les présomptions et l'hypothèse générale. Le présent projet a été axé sur le recueil et le codage de données en vue de leur intégration à des données structurelles qui seront ultérieurement utilisées pour étalonner et valider le modèle conceptuel. Un cadre descriptif a été élaboré, et des occurrences d'échecs ont été cernées au moyen de la méthodologie développée par le Projet des indicateurs-pays pour la politique étrangère (CIFP) de l'Université Carleton. Des données d'événements ont été recueillies au cours des 24 mois qui ont précédé ces occurrences d'échecs d'États. Des données d'événements ont également été recueillies au cours des périodes de stabilité relative qui ont été cernées

pour ces pays. Ces événements ont été classés selon qu'ils nuisaient ou contribuaient à l'autorité de l'État, à sa légitimité et à sa capacité, y compris la gravité du défi enregistré. Un test de fiabilité d'inter-code a été fait pour s'assurer de l'uniformité du codage. Les résultats ont été comparés avec des données obtenues de *Virtual Research Associates* (VRA)[i], ce qui a permis de comparer le codage humain avec le codage machine. Les recherches ont mené à la conclusion que le codage humain est plus discriminatoire, mais aussi beaucoup plus laborieux. Une base de données exhaustive a été créée; son analyse a été commencée et elle se prolongera après la soumission de ce rapport.

Il est prévu que le modèle conceptuel servira à intégrer des événements et des données structurelles, ce qui permettrait de mesurer et de surveiller les tensions d'États. De plus, au moyen d'une analyse de régression, il serait possible de déterminer les seuils de vulnérabilité et d'instabilité, et de prévoir les crises d'intérêt. Une base de données exhaustive unique, qui a été développée en appui à ces recherches, est documentée dans le présent rapport.

Executive summary

David Carment, Yiagadeesen Samy, Doug Hales, Jordan Miller, Liz St.Jean & Peter Tikuisis. 2010. Indicators of State Failure Phase II. CR2010-124 DRDC-Toronto

The objective of this project was to support the development of an early warning model for monitoring state instability and predicting pending failure. This complemented and extended earlier efforts to develop a conceptual model and to characterize triggers, the final events recorded between non-violence and violence (relative stability and instability). The value of a crisis early warning system of state instability is largely self evident. Timely intervention in international crises is more efficient and effective than delayed response when a series of follow-on externalities that could have been prevented must be addressed. A rigorous predictive methodology could be used to consider explicit resource allocation trade-offs and provide proactive decision support.

The immediate focus of this proof-of-concept effort was to collect and code events data and integrate it with structural data that will ultimately be used to calibrate and validate the conceptual model. A taxonomy was established and incidences of failure identified using the methodology developed by Carleton University's Country Indicators for Foreign Policy (CIFP) project. Events data were then collected for the 24 months preceding these instances of state failure. Periods of relative stability for these states were also identified and events data collected for these periods. These events were distinguished as eroding or bolstering state Authority, Legitimacy or Capacity, including the severity of the challenge recorded. An inter-coder reliability test was conducted to confirm coding consistency. The results were compared with data available through Virtual Research Associates (VRA), thus affording an opportunity to gauge the merits of human (versus machine) coding. The research effort concluded that human coding is more discriminating but also considerably more time consuming. An extensive data base has been developed and analysis commenced, which will continue beyond the submission of this report.

This report recommends a number of items to continue the breadth and quality of the research in EWMs. First, events data collection should continue to include more countries over a longer period of time to create a bigger sample upon which to draw tendencies. While the results of this report stand on their own, some limitations were noted in terms of statistical validity based on sample size. Additional events data collection should reduce these concerns. Second, the definition and interpretation of crisis of interest (COI) should be refined. The current working definition provides sufficient precision for initial research; however with additional data the definition of crisis should be refined accordingly. Third, the threshold of what constitutes relative stability and instability should be refined as a means to guide case selection more effectively. This may include emphasizing certain dimensions within the existing framework or including new ones. Last, the research findings should be broadly disseminated to generate awareness within the academic and defence communities of the work conducted and to receive critical feedback to support the ongoing development of the early warning model of state instability.

Sommaire

David Carment, Yiagadeesen Samy, Doug Hales, Jordan Miller, Liz St.Jean & Peter Tikuisis. 2010. Indicators of State Failure Phase II. CR2010-124 DRDC-Toronto

Le présent projet avait pour objectif de soutenir le développement d'un modèle de pré-alerte permettant la surveillance de l'instabilité des États et la prévision d'échecs imminents, ce qui a complémenté et prolongé les recherches antérieures sur le développement d'un modèle conceptuel et la caractérisation des déclencheurs, soit les derniers événements enregistrés entre la violence (instabilité) et la non-violence (stabilité relative). La valeur d'un système de pré-alerte de crise d'instabilité des États est plutôt évidente; l'intervention rapide dans les crises internationales est plus efficiente et efficace qu'une réaction tardive lorsqu'une série d'effets externes, qui auraient pu être évités, doivent être traités. Une méthodologie rigoureuse de prévision pourrait servir à envisager des compromis explicites en matière de répartition des ressources et à offrir de l'assistance proactive pour la prise de décision.

Ces recherches de validation se concentraient surtout sur le recueil de données d'événements et leur codage, en vue de les intégrer à des données structurelles qui serviront ultérieurement à étalonner le modèle conceptuel et à le valider. Une taxinomie a été mise en place, et des occurrences d'échec ont été cernées au moyen de la méthodologie développée par le Projet des indicateurs-pays pour la politique étrangère (CIFP) de l'Université Carleton. Des données d'événements ont été recueillies au cours des 24 mois qui ont précédé ces occurrences d'échecs d'États. Des données d'événements ont également été recueillies au cours des périodes de stabilité relative qui ont été cernées pour ces pays. Ces événements ont été classés selon qu'ils nuisaient ou contribuaient à l'autorité de l'État, à sa légitimité et à sa capacité, y compris la gravité du défi enregistré. Un test de fiabilité d'inter-code a été fait pour s'assurer de l'uniformité du codage. Les résultats ont été comparés avec des données obtenues de *Virtual Research Associates* (VRA), ce qui a permis de comparer le codage humain avec le codage machine. Les recherches ont mené à la conclusion que le codage humain est plus discriminatoire, mais aussi beaucoup plus laborieux. Une base de données exhaustive a été créée; son analyse a été commencée et elle se prolongera après la soumission de ce rapport.

Plusieurs recommandations du présent rapport visant à maintenir l'étendue et la qualité de la recherche sur les EWM. Premièrement, le recueil de données d'événements devrait continuer à englober plus de pays sur une période plus longue en vue de créer un échantillon plus gros à partir duquel déterminer des tendances. Bien que les résultats du présent rapport parlent d'eux-mêmes, quelques restrictions ont été remarquées en matière de validité statistique en raison de la taille de l'échantillon; des données d'événements supplémentaires devraient remédier à la situation. Deuxièmement, la définition et l'interprétation de « crise d'intérêt » (COI) devraient être précisées. La définition pratique actuelle est suffisamment précise pour des recherches initiales; toutefois, la définition de ce qui constitue une crise devrait être précisée en fonction de l'apport de données supplémentaires. Troisièmement, ce qui constitue le seuil entre la stabilité relative et l'instabilité devrait être précisé en vue d'une sélection de cas plus efficace, ce qui pourrait correspondre à mettre l'accent sur certaines dimensions du cadre déjà en place ou à y ajouter de nouvelles. Enfin, les découvertes de la recherche devraient être largement diffusées pour attirer l'attention du milieu de l'enseignement et des milieux militaires sur les recherches menées, et pour recevoir des commentaires en appui au développement en cours du modèle de pré-alerte d'instabilité des États.

Table of contents

List of figures

List of tables

Acknowledgements

We thank the following graduate students and research associates from the Norman Paterson School of International Affairs, Carleton University, for their outstanding research assistance; Erin Bresnahan, Milana Nikolko, Simon Langlois-Bertrand and Priyanka Debnath. Thanks also to DRDC and Carleton University for administrative support, Liz St Jean,DND-staff for feedback on earlier work, CIDA for support on previous iterations of the CIFP methodology, and David Mandel for expert opinion and support and Philip Terhaar for excellent software support on the management and analysis of events scores.

This page intentionally left blank.

1 Introduction

1.1 Background

The Canada First Defence Strategy states that "… instability and state failure in distant lands can directly affect our own security and that of our allies."[1] Thus, early warning analysis has become both a topical issue and an increasingly important area of research in academic and policy communities. The Canadian government has recognized the need for predictive analysis, and DRDC's continuing work in developing a conceptual early warning model (EWM) for state instability reflects this. The US Government has also acknowledged that this requirement is likely to persist. The recent Quarterly Defence Review suggests that "the changing international environment will continue to put pressure on the modern state system, likely increasing the frequency and severity of the challenges associated with chronically fragile states" and concludes that "threats to (US) security in the decades to come are more likely to emanate from state weakness than from state strength".[2]

Building on the prior research and insights from the Country Indicators for Foreign Policy (CIFP) project at Carleton University, the early warning model presented combines baseline structural data with dynamic events data in an effort to construct an integrated and comprehensive model of state instability, as illustrated below (Figure 1: Index model

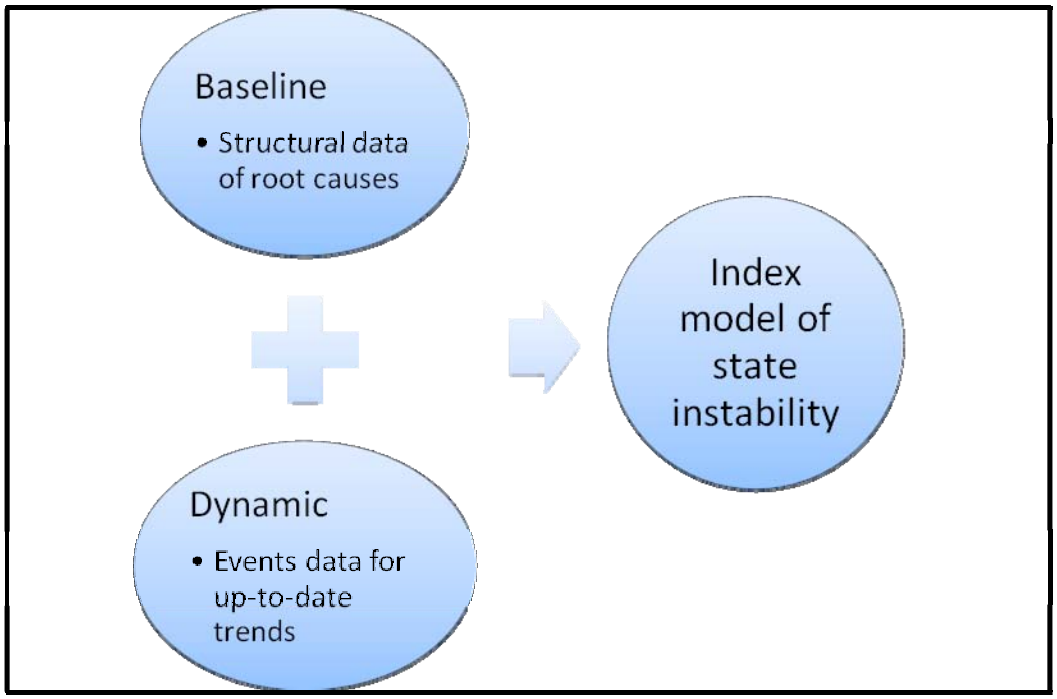

Figure 1: Index model

[1] Canada First Defence Strategy – Strategic Environment. http://www.forces.gc.ca/site/pri/first-premier/defstra/enviro-eng.asp

[2] Quadrennial Defense Review Report, US Department of Defense, February 2010 http://www.defense.gov/qdr/, pages 7 and 74.

1.2 Objectives

The goal of this project is to support development of an early warning model (EWM) for monitoring state instability and predicting pending failure. The purpose of this phase of work was to collect and code events data, and integrate them with state structural data for eventual calibration and validation of the conceptual model. Specifically, the project involved:

- Model definition,
- Identification of cases of state failure and test periods,
- Creation, comparison and synthesis of structural and events datasets,
- Econometric analysis of events data and specification for model testing, and
- Findings, Conclusion, and Recommendations.

1.3 The Team

The team consisted of CAE Professional Services (Canada) Inc., the Country Indicators for Foreign Policy (CIFP) project at Carleton University, and the Scientific Authority at DRDC Toronto. CAE brought extensive corporate experience in modelling and simulation and familiarity with the public security and public safety realms. The Carleton University team contributed academic experience and domain expertise, and DRDC Toronto, oversight and analytical proficiency. Overlap was both deliberate and constructive, providing for collaboration, challenge and mutual support. Work was shared and progress monitored on a web service (Igloo) hosted at Carleton. It was used as a common repository for source data, meeting minutes, reports in progress, and other project documentation. The team conversed regularly by email and phone, and met periodically in person for hosted workshops (17 February 2009, 22 May 2009 and 10 November 2009).

The project was exploratory in nature, essentially a matter of mapping organizational structures and human behaviour against the events portrayed by media outlets and analyzed by expert opinion. Data collection, refinement and calibration tasks proved a much larger and more complex undertaking than initially anticipated. The task of data collection was de-limited during the life of the project, with priority given to collecting sufficient data to test the EWM. Towards the end of the project, data collection and coding had been refined to a level that made collection and analysis simplified, faster and more accurate, which bodes well for future iterations of the project. This attests in part to the exemplary support received from a number of research assistants, all of whom made excellent progress in collecting and coding the source data.

1.4 The Approach

The past two decades have witnessed increasing attention being paid to fragile and failed states because of their potentially de-stabilizing impact internationally. The CIFP approach to state fragility is based on more than 17 years of experience developing a methodology and creating innovative tools for comprehensive risk assessment.

CIFP initiated development of a multi-dimensional method of analyzing country risk following award of a contract from the Department of National Defence (DND) in 1997. It drew upon DND's nascent

GEOPOL prototype as a template. This early research established the necessity of using multiple 'clusters' of data to provide a comprehensive assessment of the structural factors that affect country stability. These included, for instance, domestic conflict, political governance, and environmental stressors. Since 1997, CIFP has expanded and enhanced this approach using core funding from the Department of Foreign Affairs and International Trade (DFAIT), the Canadian International Development Agency (CIDA) and the International Development Research Centre (IDRC), in addition to a project for the Joint Research Centre of the European Commission and Petro-Canada.

The investment enabled CIFP to refine its techniques and develop training capabilities in conflict prevention and early warning analysis. As a result, the CIFP methodology has been adopted by domestic and international agencies, such as SIPRI, the World Bank, the United Nations Development Program (UNDP), the United Nations High Commissioner for Refugees (UNHCR), Brookings, and OCHA (Office for the Coordination of Humanitarian Affairs). It is particularly noteworthy that CIFP has adopted a broader understanding of state fragility through the crucial assumption that state fragility and failure can occur for different reasons. CIFP adopted an analytical approach, referred to as ALC that distinguishes three dimensions of statehood: Authority, Legitimacy and Capacity. A state with weak Authority is unable to provide a secure and stable environment and cannot enforce its laws. Poor Legitimacy refers to lack of public support for the government generally or specific policies. Lastly, a state is weak in Capacity when it cannot mobilize public resources for productive purposes.

CIFP also developed a second innovation allowing for Subject Matter Expertise and interpretation of events data. Incidents are evaluated along three dimensions (causality, escalation, and centrality), each of which is assessed according to the context of the situation. This methodology thus allows for analysts to discriminate between the effects of similar events in different contexts. In addition, this study made use of structural data from the Political Instability Task Force (PITF) and events data from Virtual Research Associates (VRA).

2 Definitions, Concepts and Terminology

Early warning and its relationship to conflict onset and state failure (and fragility) requires a solid analytical and conceptual base from which a common understanding of cause and effect can be derived. Developing an agreed, shared lexicon posed an initial challenge. To that end, the project team identified a number of terms and concepts requiring further refinement and clarification. Building on a literature review and prior experience, definitions were developed, concepts expanded and implications for the modeling effort considered.

2.1 Risk Assessment and Early Warning

Risk assessment provides policy relevant forecasting, anticipation of the likelihood of an event happening and inkling of impact and implications. [3] They are diagnostic, prescriptive, and take the form of a conditional generalization. Risk assessments precede and complement early warning - by themselves, they cannot be expected to provide precise points at which specific events are likely to occur. Risk analysis involves a calculation of an expected loss or gain associated with an event, measured by combining the magnitudes and probabilities of all of the possible negative or positive consequences of the event.[4]

In contrast, early warning usually involves combining modelling, monitoring and assessment in advance of a conflict measured in temporal terms – months or years. Late or urgent warning is the communication of imminent threat or danger. Early warning embodies a proactive political process whereby networks of organizations each with contextual knowledge, requisite skills, and information sources and capabilities, contribute analysis in a coordinated communal effort. Early warning can be thought of as the systematic collection, collation and study of different types of information coming from areas of crises for the purposes of: anticipating the escalation of conflict, developing strategic responses to these crises, and presenting options to decision makers.

Forecasting has traditionally referred to the estimation of the probability that some event will occur while the associated term gravity is used to describe the event's expected consequences.[5] Hence, the calculated risk associated with an event is the product of its probability of occurrence and the scale of its consequences. This technique produces a straightforward and intuitively appealing means to inform policymakers' risk mitigation and resource allocation management. Response plans are typically focused on events that are likely to occur and/or will be consequential.

The clear-cut distinction between early warning and risk assessment is summarized as:

> Risk assessments…identify situations in which the conditions for a particular kind of conflict…are present. They are not predictions in the sense that is usually meant by the terms "forecast" or "early warning" because risks are assessed on the basis of background and intervening conditions—the conditions that establish the potential for conflict. Whether or not risks are realized depends on whether the preconditions remain unchanged and on the occurrence of accelerating or triggering events. Early warnings by contrast are, are derived

[3] Carment, D. *Assessing State Failure: Implications for Theory and Policy"*. Third World Quarterly. Vol 24, No 3. Pp407-427.
[4] Mandel, D.R. *Violations of coherence in subjective probability: a representational and assessment processes account.* DRDC Toronto: 2008.
[5] Carment, D; Samy, Y; Prest, S. *Approaches to Country Risk Analysis and Early Warning.* Economia Internazionale. Vol 62, No 3. Pp297-323.

> from monitoring the flow of political events, with special attention to actions that are likely to precipitate the onset of conflict in high-risk situations. Risk assessments provide the context. Early warnings are interpretations that the outbreak of conflict in a high-risk situation is likely and imminent." Thus, early warnings are undertaken with the goal of presenting analyses and policy recommendations that will assist policy makers in taking action to prevent, contain and mitigate economic, humanitarian and environmental crises, and the outbreak of violent conflicts and the collapse of fragile states.[6]

The present variant of the EWM is concerned primarily with identifying antecedents to Crises of Interest (COI) - defined in detail below - which denote decisive turning points in state stability. Conceptually (and central to the EWM hypothesis) COI are linked to structural changes and significant destabilizing events that precede them. Determining the number, intensity, timing and type of preceding events and understanding relational associations that produce a COI was a core objective of this research.

An inherent challenge arises from the fact that prior causes of COIs are difficult to identify because of the plethora of potential antecedents. COIs occur not just in extreme cases of state failure but also in instances where core state structures may not be at undue risk of being destabilized. Context is of fundamental import. Characterization poses a second challenge. A state's fragility can be ascribed to non-violent factors, such as an inability of the state to deliver basic services to its population. It was understood from the onset that the selection of cases of instability should therefore be based on data from the CIFP multidimensional index and ALC framework, offering a broad and representative range of indicator inputs.

2.2 Understanding Failure, Fragility and Conflict

The start point was to develop a common understanding of how the academic and policy communities interpret state failure, fragility, conflict, and related terms, and the departure point the framework/lexicon developed by Senior Researchers at CIFP. Key descriptions included[7]:

Collapsed States: States in which the central government no longer exists. These nations exist purely as geographical expressions, lacking any characteristics of state authority, legitimacy, or capacity.

Failed States: States characterized by conflict, humanitarian crises, and economic collapse. Government authority, legitimacy, and capacity no longer extend throughout the state, but instead are limited either to specific regions or groups.

Fragile State: Fragile states lack the functional authority to provide basic security within their borders, the institutional capacity to provide basic social needs for their populations, and/or the political legitimacy to effectively represent their citizens at home and abroad.

Fragility is a process not an outcome. It is measured by a state's performance in authority, legitimacy and capacity. It is also measured by a state's willingness and its absorptive capacity and by its vulnerability to exogenous shocks

Fragmented State: Central government still functional and effective in areas under its control; unable to provide public goods to portion of its citizens in territory outside government control; conflict enduring in nature

[6] Gurr, T.R; Marshall, M.G. *"Assessing Risks of Future Ethnic Wars."* Ch 7 & App B in Gurr, TR, People versus States: Minorities at Risk in the New Century. United States Institute of Peace Press: Washington. 2000.

[7] Carment, D; Prest, S; Samy, Y. *Security, Development and the Fragile State: Bridging the Gap Between Theory and Practice.* Routledge. 2009.

Instability: a state is at risk of instability when its government faces an internal challenge in the form of a coup or civil war

Intrastate Conflict: a conflict occurring internal to a state between two or more factions or groups, or state centre and rebels, which may or may not spillover into the international domain. Intrastate conflict may occur between a government and a non-state group (internationalized intrastate conflicts often involve foreign support to one, the other, or both).

State Failure: the concept of state failure according to the PITF includes a wide range of civil conflicts, political crises, and massive violations of human rights that are typically associated with state breakdown. For the purposes of this project, the PITF identifies at least four kinds of state failure: (1) revolutionary wars, (2) ethnic wars, (3) mass killings, and (4) adverse or disruptive regime change.

Strong State: States in control of territory and boundaries, willing and able to deliver a full range of public goods to their citizens. Such states are able to withstand significant external shocks without requiring large amounts of external aid. Examples: Czech Republic, Brazil.

Tension: a measure of state status along a continuous stability/instability spectrum

Vulnerable State: a vulnerable state is one that is susceptible to exogenous shocks and lacks the capacity to deal with them; it is at risk of becoming unstable.

Unstable State: an unstable state is one that is particularly susceptible to both exogenous and internal shocks, and may be at risk of failure or collapse.

Weak States: States that are susceptible to fragility or failure because of limited governance capacity, economic stagnation, and/or an inability to ensure the security of their borders and sovereign domestic territory. A weak state is close to or at a vulnerable stage. The relationship between some of these concepts is represented in ***Figure 2*** and ***Figure 3*** below.[8] Figure 2 uses a hypothetical scaling technique to describe the four categories of state failure types. The scaling is that used in the CIFP conflict risk indexing methodology describe here: www.carleton.ca/cifp. The key features of well known types of failure, fragment, collapsed and weak states are provided. Figure 3 provides an alternative perspective on the relationship between failed and fragile states. It is clear from figure 3 that fragile states encompass all dimensions of weak, collapsed and failed states and overlaps with features found in both developing and democratising states. These dimensions are described in greater detail in Carment, Prest and Samy (2009).

[8] Carment, D; Prest, S; and Samy, Y. *Security, Development and the Fragile State: Bridging the Gap Between Theory and Policy. Routledge.* 2009.

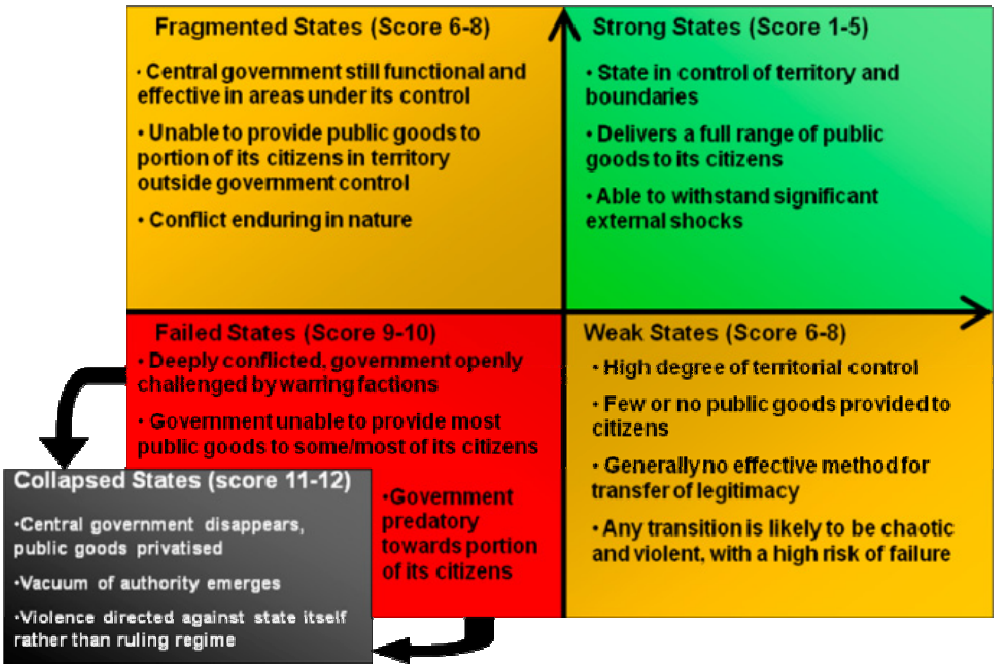

Figure 2: State Failure Quadragram[9]

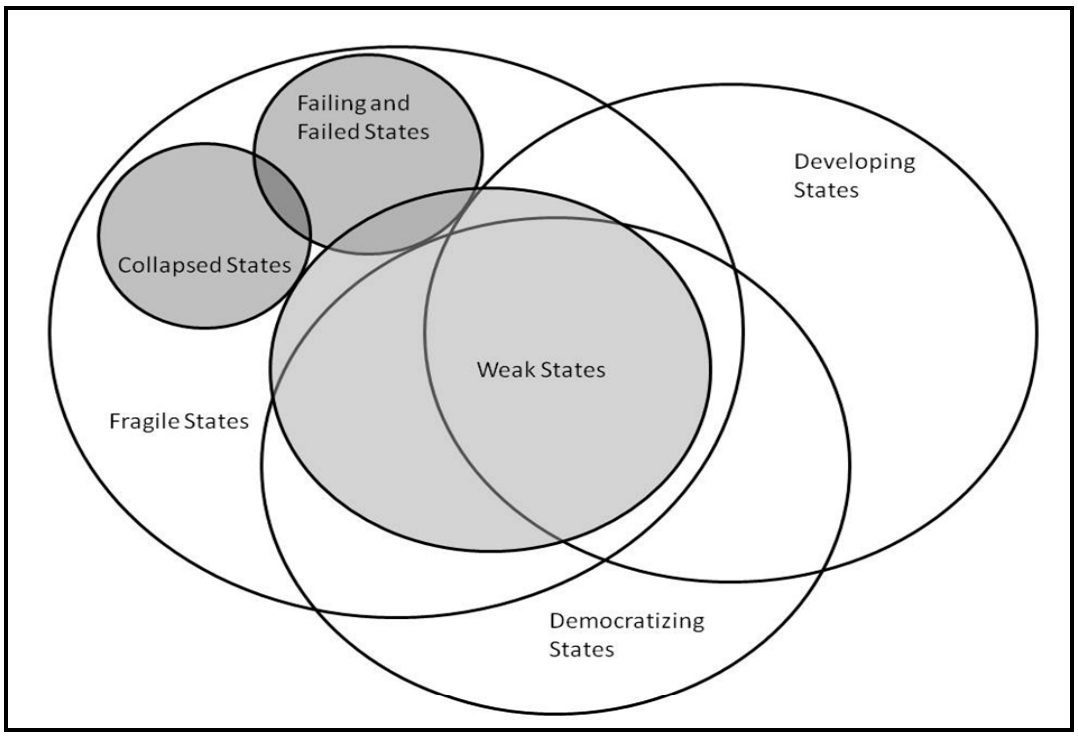

Figure 3: State Failure Venn Diagram[10]

[9] This diagram is taken from the CIFP Concept Paper 2005, available at www.carleton.ca/cifp. Copyright authors.
[10] This diagram borrows from concepts developed in Carment, D. Prest, S and Samy Y.Security, Development and the Fragile State: Bridging the Gap Between Theory and Policy (Routledge 2009). Copyright authors.

State failure and fragility are especially noteworthy as suggested in the diagrams above. Typically they result from a combination of deep underlying structural problems and sudden events-based triggers such as key leader assassinations or coups. Other terms applied to failed or fragile states include Low-Income Countries Under Stress (LICUS), poor performers, weak performers, and countries at risk of instability.[11] For its part, state failure, the overarching concept, is defined by the CIA's Political Instability Task Force as the collapse of authority of the central government to impose order in situations of civil war, revolutionary war, genocide, politicide, and adverse or disruptive regime transition. The Task Force definition weighs conflict and governance factors significantly in its analysis and, hence, its over-riding concern is with questions of instability.

Though the predominant view is that failure and conflict go hand-in-hand, there are other perspectives. For example, Robert Rotberg characterizes failed states as being distinguished by an inability to provide basic political goods, including human security and security, dispute resolution and norm regulation, essential political freedoms, and economic opportunity to most, if not all, of the population.[12] Capturing the diversity of failed state environments, Jean-Germain Gros specifies a detailed taxonomy of five different failed state types: chaotic, phantom, anaemic, captured, and aborted.[13] Their dysfunction derives from different sources, internal and external, and as a result to be successful intervention requires different (tailored) policy prescriptions. The British Department for International Development (DFID) defines state weakness in broadly similar terms, focusing on states in which the government cannot or will not deliver core functions to the majority of its people, including the poor.[14] For its part, the German Government's Action Plan on Civilian Conflict Prevention, Conflict Resolution, and Post-Conflict Peace-Building, describes failed and failing states as being characterized by a gradual collapse of state structures and a lack of good governance.[15]

The above clearly demonstrates that state failure defies simple definition and causal explanation. For example, there is an inevitable tension between the inclusiveness found in the German definition and the specificity of the Instability Task Force definition. While the latter may provide greater analytical power, the former may be of greater political utility. One trait that appears unfailingly in all of the definitions is that failed and fragile states are qualitatively different from other types of developing states, with unique problems that require innovative policy responses.

In the United Kingdom, both the DFID and the Prime Minister's Strategy Unit (SU) have released policy and strategy documents in response to a growing concern over state failure. While the DFID statement is largely policy oriented and focuses on development and aid related aspects of state failure and fragility, notably it does include a call for closer cooperation between all relevant sectors of government. The DFID document also identifies a need for improved early warning and better analysis but no specifics on the mechanics for assessing instability. From a methodological perspective, the SU's offering provides a more comprehensive framework for responding to what it refers to as "countries at risk of instability."

The SU documents (Figure 4,

[11] Ibid.
[12] Rotberg, R.When States Fail: Causes and Consequences. Princeton University Press; Princeton. 2004.
[13] Gros, J.G. *Towards a taxonomy of failed states in the New World Order: decaying Somalia, Liberia, Rwanda and Haiti.* Third World Quarterly. Vol. 17, no 3. 1996. Pp 455-471
[14] Carment, Prest, Samy. 2009
[15] Ibid.

Figure 5,
Figure ***6***) outline a detailed process intended to generate better prevention and response strategies for such states, with an assessment model that incorporates endogenous and exogenous (de)stabilising factors, country capacity, and potential shocks into the analysis of stability. The response strategy also contains a component for the identification and assessment of UK interests in intervention and the potential consequence of action or inaction. These figures have been reproduced because they provide both graphical depiction of the problem space and a useful analytical framework for situating state failure. Figure 4 provides clarity on the role of exogenous shocks in destabilising states. The SU's perspective is that the likelihood of a state entering into crisis is a function of its resilience to these shocks and its internal vulnerabilities. A feedback process exists such that the more destabilising the event is, the more likely capacity will be weakened and in turn the greater the probability that the state will be more vulnerable to future shocks. Thus destabilisation and eventual failure can result if efforts are not made to reduce internal and external vulnerabilities. Examples that clarify this relationship are provided in Figure 6. Figure 5 pertains to the time line for structural and operational response strategies. It is clear that if an external engagement is going to have a broad structural impact that a long term view is required in order to underpin key structural features of a state. As the response approximates near real time, operational strategies become more prevalent.

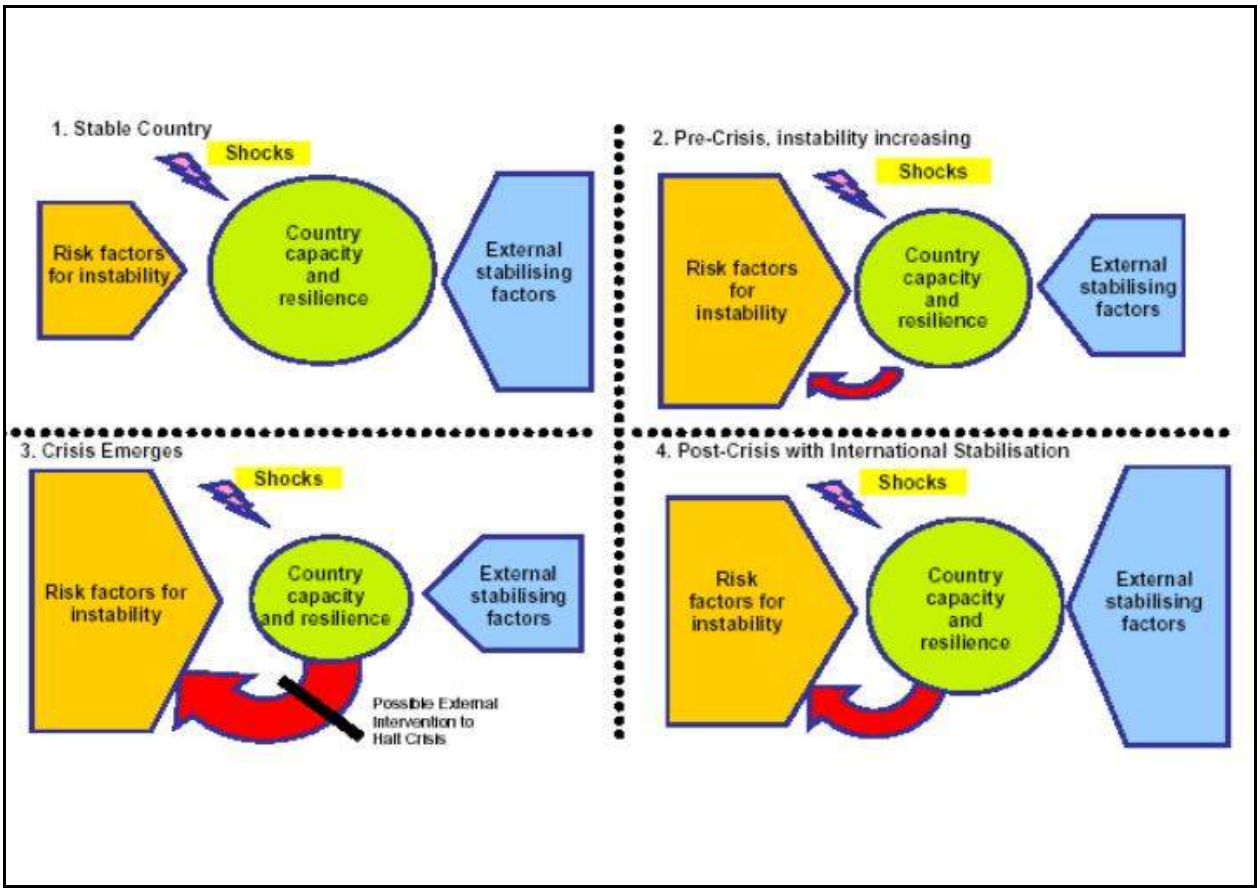

Figure 4: UK PMO Strategy Unit Analytical Framework

Figure 5: PMO Strategy Unit Analytical Framework

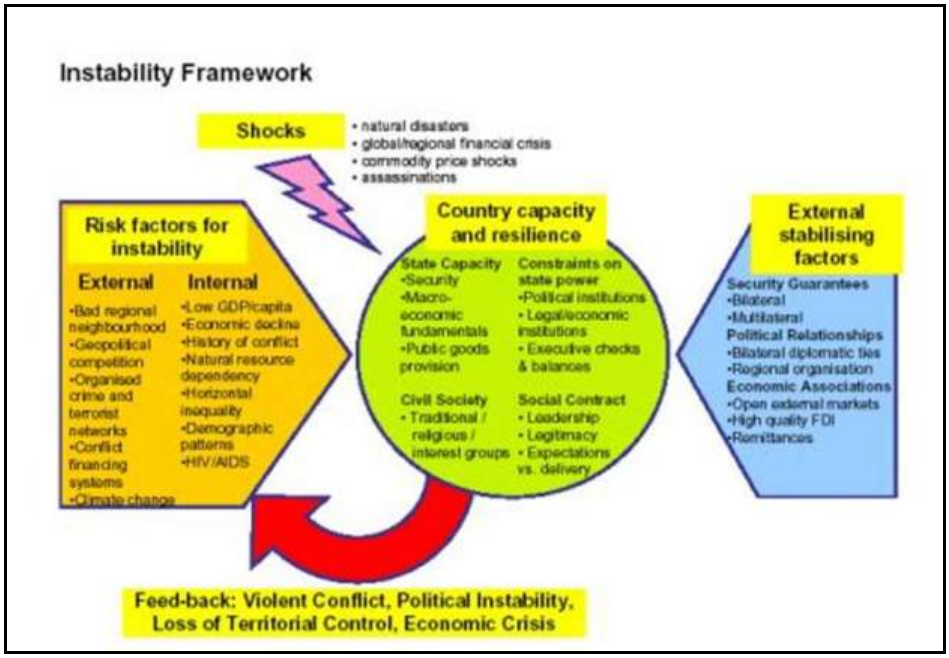

Figure 6: Instability Framework[16]

The USAID framework is depicted below (
Figure 7). While the UK's SU framework related intervention policy options to time horizons, the USAID framework identifies a generalized set of basic goals to guide government-oriented development in such regions, including: enhanced legitimacy to justify policy, sufficient will to create policy, and effective authority to implement policy.[17] To achieve such ends, the US approach echoes many of the themes found elsewhere in the current literature, including cooperation among actors at all levels, programming flexibility, and an emphasis on points of entry and leverage most likely to produce results. The diagram merely shows how USAID views types of states that are considered fragile. The key contribution here is the inclusion of legitimacy, which furthers our understanding of fragility since most studies focus only on capacity or authority. When legitimacy and effectiveness are considered low, third party engagement becomes difficult because entry points are difficult to locate and sustain. On the other hand, states falling into the upper left hand quadrant would be considered "difficult partners." These states are "strong" in some sense though they lack legitimacy. USAID sees democratic processes as a proxy for legitimacy – a viewpoint not universally shared but close to CIFP's conception of the term.

[16] Figures 4, 5 and 6 are all reproduced from the CIFP Concept Paper (2006) available at www.carleton.ca/cifp. All three diagrams pertain to work conducted by the United Kingdom Prime Minister's Strategy Unit and are documents that were provided to Carment while he worked as a consultant to their project. The Unit has since been disbanded. We are grateful to the Strategy Unit for providing these diagrams to be reproduced here. The Cabinet Office: the Prime Minister's Strategy Unit (UK). *Investing in Prevention: An International Strategy to Manage Risks of Instability and Improve Crisis Response.* A Prime Minister's Strategy Unit Report to the Government. February 2005. Pp37, 65.

[17] Fragile States Documents. USAID; Washington. 2005

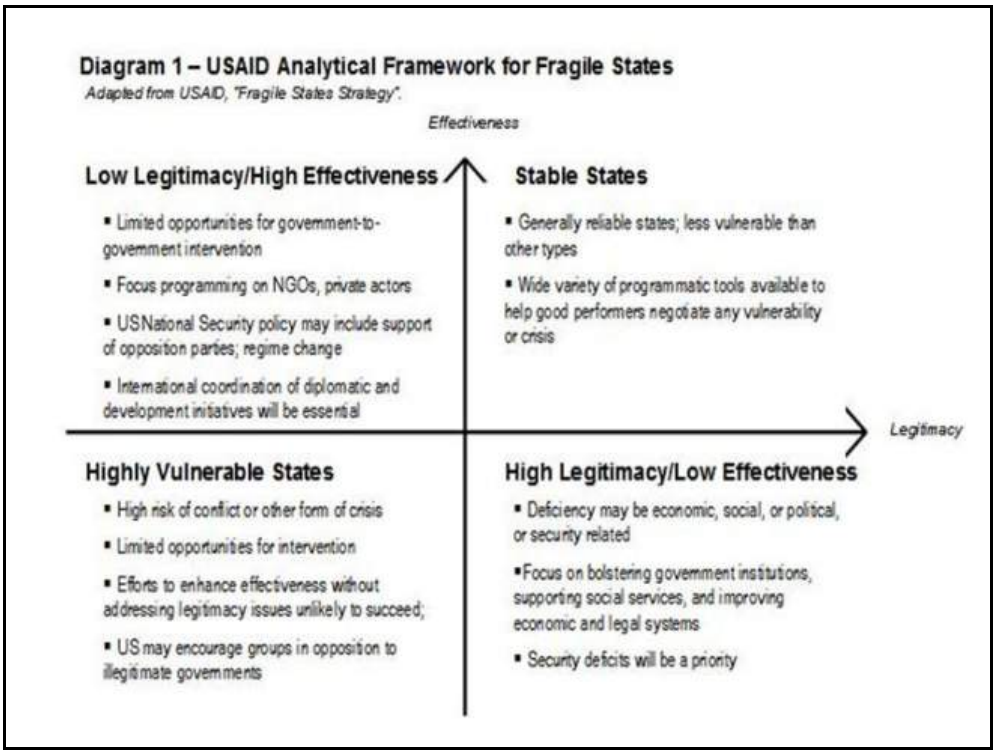

Figure 7: The USAID Analytical Framework for Fragile States[18]

Another comprehensive framework has been prepared by the Conflict Research Unit of the Netherlands Institute of Foreign Affairs (Clingendael) for the Dutch Ministry of Foreign Affairs.[19] At the core of the Clingendael methodology is the Stability Assessment Framework (SAF). The SAF integrates a number of elements into the analysis: macro-level structural indicators; institutional capacity; political actors; and policy interventions. In addition, the appraisal process incorporates a subjective component; a workshop is convened bringing together policy-makers, staff members, and local partners. This workshop is intended to offer an opportunity for dialogue, information sharing and consensus building. It serves to consolidate the stability assessment and constitutes a forum in which to promote discussion and explore policy intervention options.

An obvious strength of these assessment methodologies is their exploitation of multiple sources of data and fusion of analytical approaches. This *modus operandi* was employed earlier by the London-based Forum on Early Warning and Early Response (FEWER) working in partnership with research organisations, such as CIFP, and NGOs in the conflict prevention field. FEWER promoted a highly integrated and comprehensive framework, combining risk assessment and early warning.

An additional strength that the above analytical frameworks offer relates to the fact that each one provides an assessment of the impact that particular instances of state fragility or failure may have on national interests and an analysis of potential consequences that may follow from engagement. Building response strategies based on a foundation of relevancy enhances the likelihood that states will engage in a

[18] This diagram is reproduced in the CIFP concept paper available at www.carleton.ca/cifp.

[19] "*The Stability Framework: Designing Integrated Responses for Security, Governance and Development*". Clingendael Institute: January 2005.

sufficiently robust and sustained manner to ensure a positive and measurable impact on the incidence of state fragility or failure.

These frameworks remain works in progress. Despite multi-varied and comprehensive approaches to creating development, security, and diplomatic tools, the frameworks and supporting analysis have often proven insufficient to direct efforts and monitor progress and support the task of stabilizing and rehabilitating failed states. Given the enormous difficulties associated with programming in such environments, many governments now believe that intervention and outside involvement in an increasingly integrated world must be coordinated at the strategic level. Not surprisingly there have been and continue to be attempts to reach a level of consensus on issues of vital importance to assisting in failed and fragile states. These have generated shared insights.

One area of consensus is that policy must be grounded in an ongoing process of risk assessment and monitoring (as noted above). Supporting tools must be able to identify countries at risk of impending crisis and provide guidance as to the type of engagement required to either stave off or mitigate the consequences of the approaching storm. Further, it is generally accepted that the assessment must draw on the widest range of possible indicators of instability. To focus on a single factor such as governance or violence is to invite limited analysis of the problem and, potentially, ineffective intercession as a result. Additionally, the monitoring capability must provide sufficient early warning to allow for policy deliberation and resource mobilization, vital prerequisites of timely and effective intervention. A second area of emerging consensus is that a "whole-of-government" response is necessary to surmount the difficulties faced by failed and fragile states. Development alone cannot succeed in stabilizing a failing state any more than a military intervention can single-handedly establish an effective political infrastructure. Defence, diplomacy and development must work towards a common end and national efforts must be coordinated with other international initiatives. This speaks again to the requirement for an understanding of state fragility and predictive model.

3 Literature

One goal of this study was to help bridge the gap between the largely theoretical academic literature and policy application. Much of the existing failed state research deals in qualitative assessments and qualitative description of the contributing factors to state failure and instability. However, data collection from the previous phase indicates there has been no attempt to develop a statistical model of indicators of state failure to the level of granularity of events and mathematical precision that was undertaken for this project. A brief discussion of the purposes of developing an EWM for the policy community is an obvious departure point. It seems self evident that, if an EWM is going to be considered useful to the policy community, it needs to satisfy certain conditions i.e. an effective early warning model must be capable of:

1. Distinguishing the type of causal factors that lead to COI, thereby pinpointing points of entry for external actors;

2. Allowing the observer to understand the impact that COI have on different types of states ranging from simply fragile to fully collapsed;

3. Integrating near-real time dynamic events data with long-term structural information to counter time lags between developments on the ground and their reflection in significant indicators, and impact on programming priorities and timelines;

4. Supporting capability-based planning through sufficient and advanced long term forewarning; and

5. Providing warning of imminent conflict so that policy makers can prepare to react within days and weeks.

It follows that to establish a framework for analyzing the onset of COIs, it is necessary to understand how each given type of crisis typically develops and which possible avoidance efforts can be effective. [20] In general terms, the factors that contribute to conflict escalation are categorized as: structural factors (root causes), accelerators (precipitators/facilitators), and triggers (catalyzing events). [21] These are defined as follows:

> *Structural Factors:* Background conditions that form the pre-conditions of crisis situations such as systematic political exclusion, inherent economic inequities, lack of adequate and responsive institutions, the presence of ethnic minorities, resource exhaustion, and over-dependence on international trade.
>
> *Accelerators:* Feedback events that rapidly increase the level of significance of the most volatile of the general conditions, but may also signify system breakdown or basic changes in political causality.

[20] Gurr, T; Harff, B. "*Early Warning of Communal Conflicts and Genocide: Linking Empirical Research to International Responses"*. Tokyo: United Nations University. 1996.
[21] *A Manual for Early Warning and Early Response.* Forum on Early Warning and Early Response (FEWER): London. 1998.

Triggers: Sudden events that act as catalysts igniting a crisis or conflict, such as the assassination of a leader, election fraud, a political scandal.

Early warning theorists generally focus monitoring and analytical attention on high risk structural conditions before they fully develop, and most models focus on identifying the combination of structural factors that create specific crises. There are several reasons for this. Intuitively, one of the most obvious means of predicting future international disturbances is to extrapolate structural pattern development of the past. For example, analysts who have tracked patterns of ethnic conflict agree that regional and global observations from 1945 to the present establish that ethnically-based rebellions, much less genocide, do not just erupt spontaneously without prior indication normally extending back over many years.[22] At least in principle, such structural patterns in combination (e.g., inequality, demographic stress, neighboring state influence and conflict history) provide an origin for testing predictions regarding the correlates of crisis situations; a series of situations suspected to be of high risk can be compared to the incidence of crisis actually arising in these situations. Therefore, propositions on structural pre-conditions that are empirically well-grounded can form the basis of predictive models of events that focus on the existence of certain preconditions.[23] This is the standard approach in use by most forecasting state failure models (e.g., the PITF) using structural data (one that was not attempted here and which should be explored at a later stage).

However there are a variety of forecasting models available that go well beyond simple structure-based approaches.[24] Gupta classifies these techniques as either data-based or judgment-based. Data-based techniques involve the collection and analysis of large data sets, and judgment-based techniques involve the subjective assessment of situations by experts. Both have their strengths and weaknesses. The data model is time and effort intensive, though it does produce assessments that are clearly measureable with evidence. The judgment technique is less time intensive and captures a broader qualitative significance, though it does not have the same replicability and quantitative aspects of the data-based technique. See ***Figure 8*** below for a visual representation of both methods.

.

[22] Gurr, TR; Moore, WH. *Ethnopolitical Rebellion: A Cross-Sectional Analysis of the 1980s with Risk Assessments for the 1990s.* American Journal of Political Science. Vol 41, No 4. October 1997. Pp 1079-1103.
[23] Gurr, TR; Harff, B. 2006.
[24] Gupta, Dipak. *"An Early Warning About Forecasts: Oracle to Academics"*, in: Schmeidel & Alderman: Synergy in Early Warning Conference Proceedings. Toronto: March 15-18, 1997. Pp 375-396

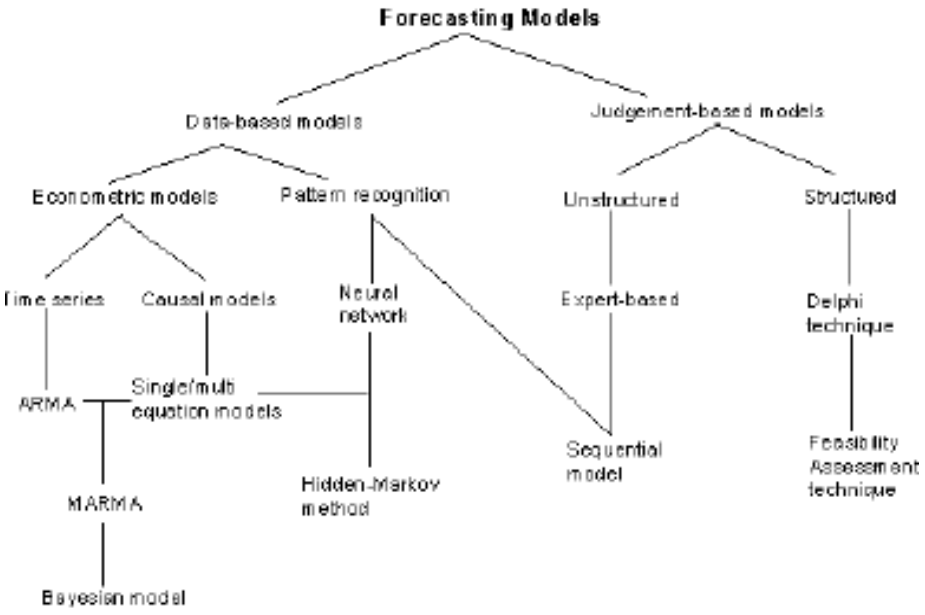

Figure 8: Forecasting Model[25]

The PITF (established in 1994 as the State Failure Task Force) identified the factors that statistically distinguished states that failed from those that averted failure over the last 40 years. The *State Failure Task Force Report: Phase II Findings* isolated three structural variables that were significantly correlated with subsequent state failure: infant mortality, openness to international trade, and level of democracy. Three separate analytical techniques confirmed these findings: logistic regression and neural network analysis estimated the predictive accuracy of the model and genetic algorithm modeling was used to identify candidate sets of variables and serve as a check on the univariate regression methodology. These findings informed subsequent research and modelling. In support of these findings but using a different data set Carment, Prest and Samy's 2008a, 2008b, and 2009 studies identified four robust structural determinants of state fragility including infant mortality, GDP per capita, growth and the level of democracy.

These structural frameworks of state failure have proved useful from a modelling perspective, but less so in forecasting outcomes for individual countries. Hence Barbara Harff developed a sequential model for early warning of genocides and politicides, resembling a process model, but incorporating accelerators and triggers.[26] She distinguished ten background conditions, four intervening conditions, and eight accelerators. What is unique about her work is that she does not assume that crisis development is linear and dependent on structural changes. Structural models, absent accelerators and triggers, identify stages of a conflict but are inherently stagnant and cannot provide for adequate risk assessment to support operational planning in response to cater to "impending" situations. Dynamics accelerator and trigger events can potentially serve as key leading performance indicators for this purpose.

[25] This diagram is reproduced from Carment, D (2003) Anticipating State Failure.

[26] Harff, B. *Early warning of potential genocide: the cases of Rwanda, Burundi, Bosnia and Abkhazia.* In Gurr & Harff (eds) Early Warning of Communal Conflicts and Genocides: Linking Empirical Research to International Responses. United Nations Press; Tokyo. pp 47-48

In another interesting and related example, Moore and Gurr used the 1991-95 data from the Minorities at Risk project to contrast and compare three empirical approaches to generating risk assessments:[27]

1. Risk profile: a list of high risk factors are generated based on general theoretical knowledge such as group incentives, capacity, and opportunity;

2. Theoretical regression model: an argument is expressed as a multiple equation model, and a statistical technique—three-stage least squares—is applied to the data to estimate the parameters of a predictive equation;

3. Empirical regression model: an inductive approach similar to the State Failure project, where statistical software determines what variables enter the analysis.

Each model yields slightly different results, albeit with a proportion of overlap. The conclusion of their study is that it is difficult to recommend one model over another; multiple approaches using different data sources should be encouraged and forecasts generated early and often. They advise that it may be useful to focus analysis on those cases that appear on multiple assessment lists. This research formed the backdrop to the current study which focused on the question of how one might combine data into one model.

Previous research,[28] suggested that there are several ways to combine events with structure at least in theoretical terms. A complete model would require an understanding and integration of:

Macro or long-term processes associated with system-structure transformations and the associated problems of the emergence of weak states;

Intermediate mechanisms associated with institutional viability and state weakness; and

Micro or short term selection processes and mechanisms that account for preferences of violence over pacific forms of strategic interactions and the subsequent escalation and/or duration of ethnic hatreds, violence, repression, and war at specific points in time.

The different types of data, and their contribution to predicting state failure is represented below in Figure ***9*** for Indexing the Model and Figure 10 for the timeline.

[27] Moore, W; Gurr, TR. *"Assessing risks of ethnopolitical rebellion in the year 2000: three empirical approaches'* in Schmeidl & Adelman. Synergy in Early Warning Conference Proceedings. September 1998.
[28] Carment, 2003.

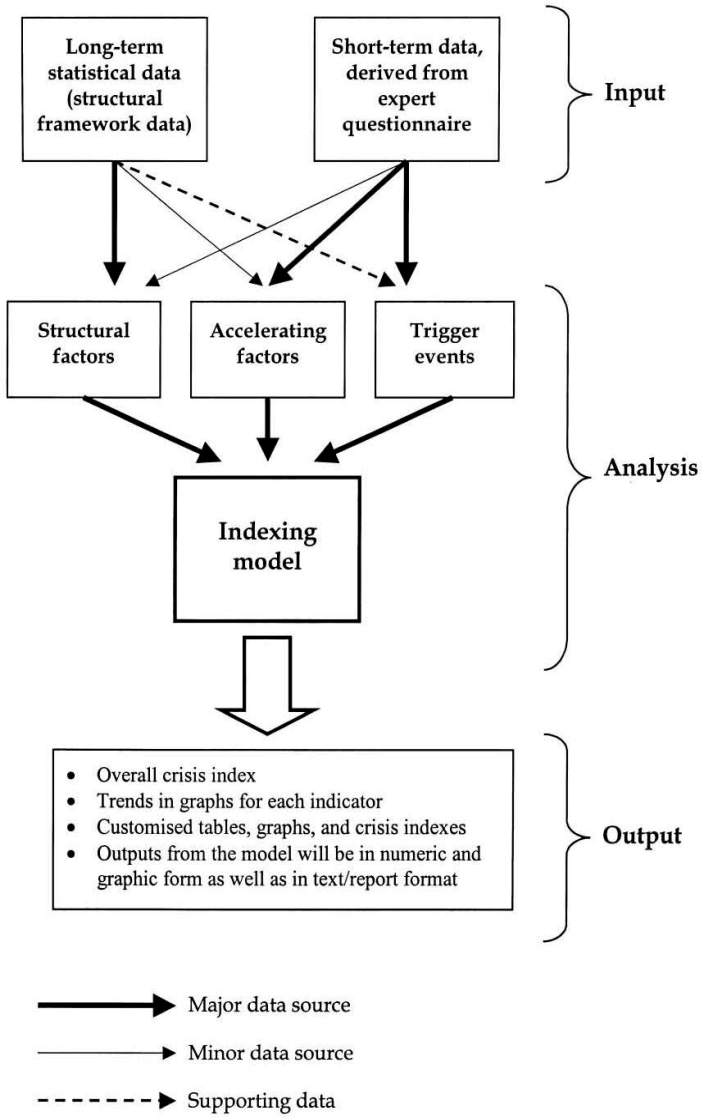

Figure 9: Types of Data for Indexing the Model - Hagmeyer-Gaverus and Weismann 2003[29]

[29] This diagram is reproduced from an unpublished paper based on a collaborative effort between CIFP and SIPRI undertaken in 2003 in order to produce an indexing model for predicting conflict using CIFP and SIPRI data. Hagmeyer-Gaverus, G. and Weismann, M (2003) Early Warning Indicators for Preventive Policy (SIPRI Working paper). Available at: www.carleton.ca/cifp.

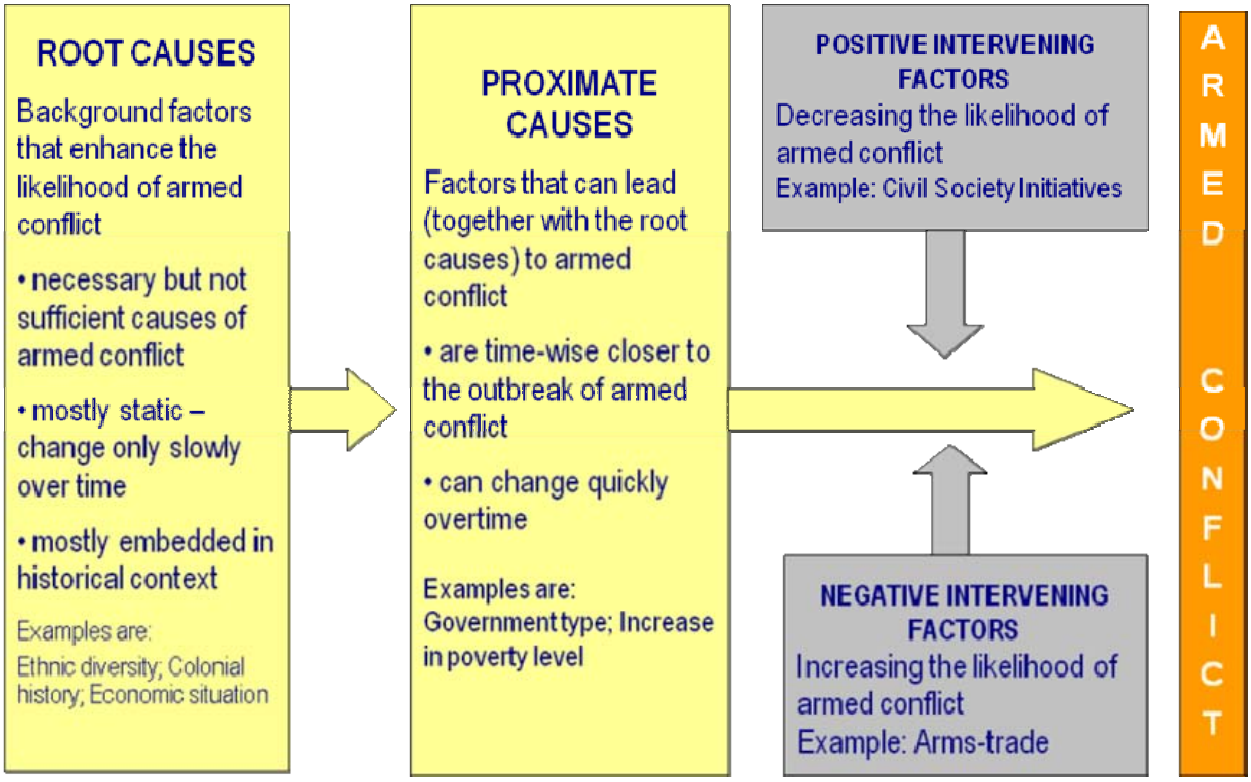

Figure 10: FAST Model Timeline[30]

An essential difference between the two representations relates to implied determinism. The focus in

Figure ***9*** is on data integration, the relationships between structural indicators and events data. The focus in the FAST Model (Figure 10) is temporal integration and sequencing – e.g. how causes relate to outcome in terms of the time at which they occur (hence the idea of root versus proximate causes). This distinction is significant and more than just conceptual; the implication is that some structures in the FAST model might be considered proximate causes. However, in the SIPRI framework these are kept distinct, an approach reflected in and more closely approximating the EWM.

Work on baseline (structural) data - exemplified by the Political Instability Task Force - has often focused on variables related specifically to overt conflict. However, work by CIFP has demonstrated the utility of adopting a broader concept of the state.[31] In the literature review and in framing the EWM, a variety of different sources of information were drawn upon and note taken of diverse levels of engagement ranging from the macro to the micro:

1a) **Macro Level evaluation** of structural indicators (econometrically or through pattern recognition techniques) (e.g. parts of the State Failure Project; PIOOM (Dutch Acronym – Interdisciplinary Research Program on Root Causes of Human Rights Violations); CIFP (Country Indicators for Foreign Policy); HEWS (Humanitarian Early Warning

[30] The FAST model is fully explained in Krummenacher et al (2006) and in various documents prepared by the Forum on Early Warning and Early Response (see ww.carleton.ca/cifp).
[31] Ibid.

Service); ICB (Institute of Canadian Bankers); FIRST, Rummel's Democide data-base, Uppsala's Conflict data-base);

1b) Macro Level time series of leading indicators (e.g. IOM; Refworld; FAO's GIEWS; Reliefweb; the UN system-wide Earthwatch; HazardNet for disasters; the global early warning system for displaced persons - GEWS);

2a) Intermediate Level conjunctural models that track changes in pre-specified events and interactions between groups (e.g. conflict/cooperation, genocide, non-violent protest) using machine-coded data, pattern recognition and neural networks (e.g.Protocol for the Assessment of Non-Violent Direct Action [PANDA]; Kansas Event Data System [KEDS]); and

2b) Intermediate Level structured (Delphi) and subjective models, which utilize a team of experts who identify key actors and estimate their future position on a given issue (regime stability, turmoil likelihood, investment restrictions and trade restrictions) with regards to their power to influence the outcome, the importance (salience) they attach to the issue, and the certainty or firmness of the actor's orientation (eg, Decision Insights; Political Risk Services).[32]

The second and most challenging aspect in the development of the EWM is that of capturing and integrating dynamic events data. Events data are important to include because they offer information and insight into the current, prevailing situation and inclination in a country. Event trends can discriminate whether a structurally problematic condition is worsening and the state is becoming increasingly unstable or whether circumstances are improving and the state is becoming more stable. If interpreted correctly, events data provide an initial indication of information that will appear in the structural data a year or more later. Events data is quintessentially micro in nature:

3a) **Micro Level** sequential models which develop risk assessments based on tracking of specific behaviours - using accelerators (e.g. parts of State Failure; CEWS);

3b) **Micro Level** response models which evaluate outside response to conflict and develop feasibility assessments based therein (e.g. Helen Fein's Life Integrity Violations Approach; the International Development Research Centre's PCIA); and

3c) **Micro Level** field reporting by NGO networks (e.g. FEWER; FAST; ICG, CIPDD) using structured and/or unstructured reporting techniques.[33]

Together, structural and events data provide a *potential predictive capacity* to identify early warnings for vulnerable states, where events may cause instability in a state. It might also offer urgent warnings for unstable states that are at risk of failure. Moreover, it might help identify formerly vulnerable and unstable states that are moving towards stability.

Building a Model: Data Inputs

Having established that problems such as conflict risk potential, failure and fragility are best understood from a multiplicity of perspectives, the challenge was to appreciate and integrate data inputs. Each approach has biases and strengths, and focuses on micro or macro views as well as on static conditions or emergent events. A survey allows the viewer to appreciate the uniqueness and the similarities. Used together, different perspectives provide a more complete and balanced picture than any one perspective. In the following section, different information streams are identified that can act as a system of checks

[32] Ibid.
[33] Ibid.

and balances on each other, with each source providing a method of validation for the others. Together they provide a more complete picture of country performance.

3.1 Structural Data

Structural data, such as GDP per capita, political indices and human rights measures, provide a sound basis on which to build country profiles. Structural data are compiled by recognized organizations, sometimes in partnership with host nations. Structural data allows the end user to rank countries for quick assessments of performance within sub-sectors. Country level structural data also enable comparative analysis. Many statistical indicators are composite indices capturing several underlying concepts in a single score – the UNDP's Human Development Index (HDI) is an example of such a composite index. Indexing makes quantitative data easier to handle and compare, and is useful for broad strategic evaluation across countries. For example, the CIFP indexing approach utilises a three-step process of initially collecting data on a yearly basis, assigning raw scores for a global rank based upon a continuous distribution of countries for each indicator and then ranking countries for a specific year.

Structural data have obvious merit from a macro or strategic perspective but, a number of factors limit their utility as the sole source of information in decision-making. At the sub-national level, variations in both the types and method of data collection tend to limit an end-user's ability to compare indicators across sub-regions or within a single region over time. In particular, sub-national data is often not delineated by age or gender, thus limiting the extent to which it can inform targeted development programming. Even at the national level, in some cases statistical data for some indicators simply do not exist or is uneven in its coverage. However, recent efforts by the World Bank and the UNDP have improved country-level data collection and reliability. Beyond these issues, operationalizing measures of conflict and failure are a challenge for country-level structural data analysis. Analysts must use specific and narrowly defined kinds of information to proxy or otherwise represent the more abstract concepts that lay at the heart of country performance.

There is a need to balance too many and too few indicators to explore underlying constructs. Too few indicators provide an incomplete picture. Too many make it difficult to distinguish vital information amid all the background 'noise'. In circumstances where structural data are unreliable or proxy measures cannot be properly identified, alternative information sources should be sought.

Data are only meaningful if they are considered in context. Context might include a comparison between neighbouring villages, between states with similar economic development or democratic history, or of the same unit of analysis over time. Country performance indices, such as that compiled by CIFP, can provide a good source of information for cross-country comparisons. Analysts should be sure to familiarize themselves with how categories are operationalized, and what the statistics mean.

3.2 Dynamic Events Data

The systematic collection and evaluation of dynamic data also known as events-based information analysis, is highly relevant to risk analysis and early warning. *Dynamic* is used deliberately and refers to the continuous change, activity, and transformation in occurrence and intensity of proceedings. Data analysis, whether it draws on information from media sources or subject matter country experts, is useful for identifying indications and evaluating emergent variations in popular perceptions, preferences and

stakeholder behaviours. Dynamic data analysis can add considerable value through regularized and standardized reporting. It can deepen an understanding of trends depicted in structural data, and reveal trend reversals. For example, a statistical study may show a steady decline in violent events over a series of years, but current events may uncover a sudden surge in violent demonstrations, one that will show up in structural data only long after the fact. Events-based information can also provide a window into stakeholder perceptions and insight into future behaviour, how they are reacting to real-time changes and why they are doing so.

The choice of policy responses to state fragility and failure is informed by the explanations used to account for their onset, decay, collapse and recovery. Most importantly, to be policy relevant those responses must cater to the needs of decision makers whose choices are constrained within fairly narrow time frames and windows of opportunity. Decision support must complement existing processes and acknowledge decision cycles, and analysis must recognize the capabilities and resources that a decision maker has at their disposal. Providing active real time monitoring in conjunction with structural data can help ensure those resources and capabilities are exploited more effectively and allocated in a timely manner.

Further dynamic data analysis is intended to identify not only the trajectory of fragile states, but also to highlight any sectors that are particularly vulnerable and any factors, either exogenous or endogenous, that may contribute to a potential for state instability and collapse. These high risk areas constitute potential entry points for intervening action. Response options can then be evaluated and formulated on the basis of their impact on elements that represent a significant destabilising or stabilising influence. In this way, those elements that pose the greatest risk and richest opportunity - and thus contribute greatly to the potential weakening or strengthening of a country - can receive priority attention. Ascertaining and addressing these risk areas offers the most expedient and efficient method for reducing a country's risk of instability.

Events data analysis has evolved over several decades. Initial examination was led by the *International Crisis Behaviour* and the *Correlates of War* projects in the 1960s and research continues today. Much of the current effort on events analysis used in conflict monitoring and early warning research is driven by the *World Event - Interaction Survey* (WEIS) methodology developed in the 1960s and since modified and updated by Joshua Goldstein, among others, through the introduction of standardized scaling. The WEIS methodology focuses on discrete events, particularly crises, and supports some of the best known studies of conflict and cooperation interactions between states. The unit of analysis in the dataset is event-interaction, referring to words and deeds communicated between nations, such as threats of military force. Each event-interaction documents a daily report of an international event.

The WEIS methodology entails assigning a numeric value based on the type of event occurring. It includes a coding of event-types assessed initially by a panel of international relations specialists and evaluated in the context of inter-state relations. All subsequent events are coded automatically using a 0-10 -10 to +8.3 scale representing the most conflcitual and most cooperative interaction respectively.[34] The distinct advantage of scoring an event based on category is the considerable ease of interpretation and standardization afforded. One can draw heavily on open source information such as newswire reports. However, analysis that relies on solely on such static interpretations of event types is at risk of producing inaccurate conclusions, given that the context of the situation matters.

[34] Veen, Tim. *Event Data: A method for analyzing political behavior in the EU*. The Methods & Data Institute, Centre for International Crisis Management and Conflict Resolution: University of Nottingham. September 2008. Pp12

There are other significant limitations in this approach. The WEIS and methodologies like it do not automatically or necessarily capture sub-state processes. To address this deficiency without jeopardizing WEIS-related benefits and insights, a number of scholars led by Philip Schrodt and Deborah Gerner have produced a reliable means of machine coding capable of deriving meaningful content from news services for the purposes of sub-state analysis. Their project known as the Kansas Events Data System (KEDS) has raised text processing and sparse parsing of language to a high level of consistency and efficiency. KEDS exploits automated coding of English-language news reports to generate political event data focusing on the Middle East, Balkans, and West Africa. These data are used in statistical early warning models to predict political change.

Using a methodology similar to that developed by Schrodt et al., the *Early Recognition of Tensions and Fact Finding* (FAST) research group led by Heinz Krummenacher, formerly of the Swisspeace Institute in Berne, used a combination of human coding and a software programme developed by the Virtual Research Associates (VRA). VRA is based in Boston and affiliated with the successful *Protocol for the Assessment of Nonviolent Direct Action* (PANDA) at Harvard which was led by Doug Bond.[35] Like KEDS and WEIS before it, PANDA was based on research intended to guide and inform the automated coding of events (news reports), both violent and otherwise. VRA has since established an *Integrated Data for Events Analysis* (IDEA) protocol currently in use by various US government agencies and regional organizations such as the Intergovernmental Authority for Development (IGAD). The IDEA protocol expands the PANDA protocol into a more generic framework or ontology suitable for use in monitoring events in social, economic and political sectors.

The FAST and VRA methodologies merit description in some detail. FAST's approach reflected an attempt to reconcile and integrate qualitative assessments and quantitative methods. Initially, open source events were automatically coded and assigned a numeric value based on its cooperative or conflictive character. These coded events were then aggregated on a weekly and monthly basis to establish an overall conflict potential and measure of the carrying capacity of a state (i.e. the ability of a state to absorb internal challenges and shocks without exhibiting significant indications of instability). The FAST approach was premised on the inadequacy of automatic coding information, not the accuracy of automated coding, so it retained features in terms of text parsing and content analysis found in the earlier KEDS and WEIS protocols. The chief criticism of earlier approaches is that, like all machine-coding approaches, they relied on just a few mainstream news wires such as Reuters for their raw data. The challenge is to ensure consistent coverage. Attention varies and a particular newswire's focus and the associated lack of attention to daily developments within a country not judged to present a risk to global security poses a problem, particularly during peaceful times.

As a result, FAST reporting mechanisms moved beyond newswire services and incorporated "Local Information Networks" (LINs), which were essentially trained field-monitors charged with logging relevant information and coding using the same rule set. Kosovo serves as an example. Using LINs, FAST was able to attain an average of 68 events per month, a higher rate of reported events than that obtained through automated coding.

The VRA - IDEA methodology is similar to FAST's but it relies, not just on field monitors who are potentially available to help gather information on emerging conflict, but also "participant-observers" as well. Doug Bond purports that the matching of participant-observers with paid informants helps create greater authenticity by taking advantage of the sense of urgency faced by the stakeholders who are most immediately affected by escalating violence. In this regard, Harvard's Global Negotiation Project applies the IDEA methodology through "The Third Side". The Third Side looks at conflicts not just from one

[35] Bond, DJ; Bond C; Oh, JCJ; Taylor CL. *"Integrated Data for Events Analysis (IDEA): An Event Typology for Automated Events Data Development."* Journal of Peace Research. Vol 40, No 6. Pp 733-45

side or the other but also from the broader viewpoint of the surrounding community i.e. it incorporates a stakeholders' perspective. Asserted stakeholder perspectives can serve as early warning signals. The result is production of a series of field reports that purportedly reflect the true nature of the evolving security environment.

FAST, VRA and related machine-coded event specific applications entail four types of difficulty. First, as noted, events data draw from a myriad of open sources assembled and catalogued by humans or via computer and through machine-coded language. In either case, when each discrete event is analyzed in a structured and systematic fashion, patterns of performance begin to emerge - but only if there is sufficient events related information collected on a sufficiently regular basis over a sufficiently long period of time. Obviously such commitments require significant resource investment. Pattern recognition is especially critical to the analyst who is engaged in continuous monitoring, whose goal is to derive projections about short term changes within a country on the basis of recent trends, and to determine in advance if a situation is deteriorating or improving. Events-based analysis cannot supply the full context for complex situations, nor can it provide a complete representation of root causes. The challenge lies in sorting through events *ex post facto* to determine which class of event might serve as antecedent to a crisis in the future. Establishing situational dependence is problematic. Not all events are significant in the sense that they may or may not be causally related to fragility or failure. For example, some events represent specific, known and standardised interactions among key players that occur routinely over a period of time. Other discrete events act as accelerants (factors that tend to magnify the effects of existing issues) either on their own or in conjunction with a series of similar events. Knowing exactly what kind of event constitutes an accelerant and what is habitual behaviour is difficult to discern in advance. Barbara Harff's sequential model for early warning of genocides and politicides, and discussed earlier in section 3, attempts to address this problem.

Discrete events can act as accelerants (factors that tend to magnify the effects of existing issues) either on their own or in juxtaposition with a series of similar events. Events can also be triggers that provide the 'catalyst' cause for crises. Events precipitate reactions and, if appropriate pre-conditions are in place, can become the basis for wholesale transformation. Establishing and modelling cause and effect is challenging. Appreciating how specific events may trigger reactions and provide the appropriate pre-conditions in advance is complicated. The assassination of Juvenal Habyarimana, the President of Rwanda, serves as an extreme but illustrative example. His assassination set off the ensuing genocide in Rwanda but, absent the underlying structural tensions and the accelerators deriving from ethnic fragmentation, power imbalances, and land shortages among other things that radical elements could exploit, such a severe reaction probably would not have occurred. In other words, his assassination served as the trigger, but under very specific enabling conditions. In this sense an analogy can be drawn to chaos theory in which system dynamics are highly sensitive to initial conditions and the objective is to establish an underlying order in seemingly random data. In essence, it is very difficult to generalize across cases when it comes to identifying triggers. While it may be possible to identify in advance the range of factors that can serve as accelerators with an impending crisis, trigger identification is much more difficult. Simply put, triggers may be *sui generis*; certainly from the perspective of analysis they must be treated as such.

These two insights led the project team to create a working definition of COIs based on types of human interaction that are known to generate instability. We recognise this definition was arbitrary in some respects and is in need of further refinement since it was developed in the absence of empirical data and is a theoretical construct. It remains to be determined how much empirical support there is for our COI definition.

As previously suggested there are additional challenges associated with machine-coding of events data. Most events data collection methodologies assign a numerical score to represent how cooperative or conflictual an event is deemed to be. The grading is determined using static interpretation and coding rules. This has an attendant problem of de-contextualization that occurs in the absence of human coding insofar as the cooperation-conflict continuum assumes universality. Significantly different types of events may automatically be treated as equals when data are collected and aggregated. The net result can be "strange equivalences"; two events may incorrectly appear to have equal impact because they are of the same type or assigned the same numerical value. More specifically, scoring is associated with a "type" of event; it is not mindful of the regional, societal or conflict-specific nature of the incident. Analysts need to look at events in the context of past history and social realities in an effort to try to understand what is driving an event. This insight led us to develop a context driven events coding procedure for this project, which is outlined below.

The third and fourth issues are, respectively, the question of which open source information to use (including local language sources) and how to account for multiple descriptions of a single incident i.e. cases when the same event is reported in different stories and newswire services. Seeking out different sources for a broader set of views can reduce, though not eliminate, the bias that might occur were one to rely on a single media source for all information. At the same time, aggregate event scaling such as the unmodified WEIS continuum has the potential for lost information and for events to be coded more than once. Some sense of context can be derived from events data collection. Given an absence of opinion polls on matters relevant to fragility, local media reports or other sources provide insight into popular reaction to events and serve as a key enabler to understand the reality of fragility on the ground and discover how a government is responding to social, cultural, and economic pressures and opportunities. Local language media are important sources of information, but relying solely on automated coding of such sources might prove unreliable. All reports of an event will most likely provide useful information and ideas but, with respect to local, and potentially partisan sources, an analyst must decide what "fact" is, what "conjecture" is, what "polemic" is, and how reliable event coverage is as a result. Analysts need to look at events in the context of past history and social realities, to try to understand what is driving an event; they must adopt a critical eye when evaluating credibility. Nonetheless, they can use local language media to help contextualize events and support sub-state monitoring.

These insights led our team to draw on a broad range of information sources rather than one newswire service and it also creates opportunities for language parsing beyond English (Arabic, Chinese, Spanish and French are just some of the languages that have been interpreted by SME event coders).

The use of field monitors to systematically record events data presents an alternative if not complement to machine coding. It offers both distinct advantages and disadvantages. Personal insight, when informed by contextual knowledge, can supply invaluably nuanced views of country fragility, though it too may have an implicit or explicit bias. Qualitative information, when systematically recorded and assessed, is an important balance to the systematic collection of newswire data, as it uncovers details and nuance. Put simply, when correctly structured, stakeholder and field monitor analysis can provide the "why" behind the "what" revealed through automated analysis. Human coding exploiting stakeholder opinion can also perform a valuable challenge function to more quantitative analysis. Perception and premonition have some inherent value. If enough stakeholders tell a story that differs from a statistical snapshot it is probably worthwhile reconsidering the validity of a quantitative-based conclusion, and potentially revising the selection and/or operationalization of quantitative indicators. Context also provides analysts with a way to infer carrying capacity. Coding is interpreted through standardised meaning and procedures that allow equivalency of events across cases but measured against the import of that event within a particular case.

However, there are some drawbacks to relying exclusively on field monitors. Individual opinion, even if systematically recorded, tends to reveal only one segment of a larger picture. People have subjective viewpoints, whether they are ordinary citizens or specialists. Specialists are likely to overestimate the importance of their particular field in evaluating the situation in a country. Ideology can cloud opinions, as can personal experience and bias. Even stakeholder or "Third Side" options cannot provide an objectively true description of a country's fragility processes. Field monitoring has other limitations as well. It is comparatively expensive relative to other methods; trained local experts require compensation, and it can be costly to train and retain a sufficient number to validate the information collected. Gathering stakeholder input takes time and many fulfill demanding and vital positions within their countries' government and society. They may sacrifice their own work to accommodate requests or be unavailable for interviews and data elicitation during a period of interest.

Typically events data draws from a myriad of open sources collected by humans or machines. In either case, when each discrete event is analyzed in a structured and systematic fashion, ordered trends begin to emerge. Pattern recognition is especially important to the analyst who is engaged in continuous country monitoring and, who wants to make projections about short-term changes within a country on the basis of recent events and system behaviour. For example, Figure 11 below depicts patterns of decline in governance performance approximately a year prior to the declared state of emergency in Pakistan in 2007. The red (lower) regression line in the graph portrays the overall "events" trend. Clearly, in this case, there was considerable suggestion of an approaching crisis. Such indication, if properly understood, can allow policymakers to respond in a timely fashion to impending problems, rather than simply responding after the fact i.e. to be more proactive and less reactive.

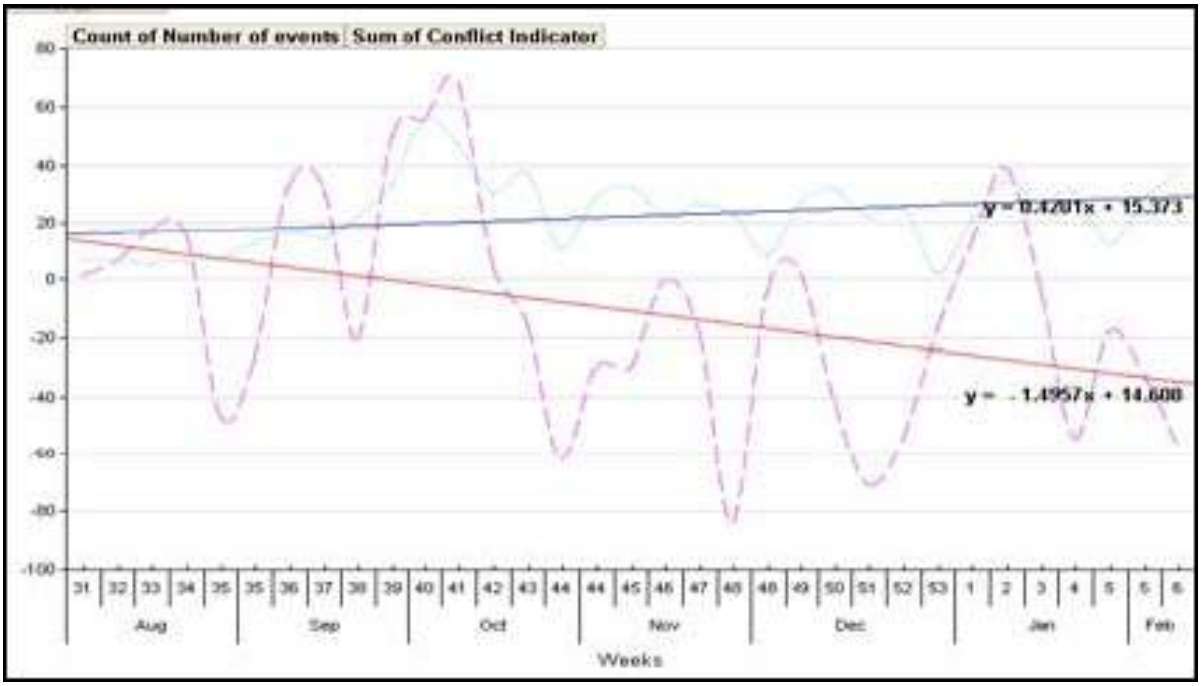

Figure 11: Pakistan Governance Events[36]

[36] Diagram reproduced from the CIFP Pakistan Democracy and Governance Report (2007) located at www.carleton/cifp.

In sum the project undertook to investigate how to collect and characterize dynamic events data in order to integrate this with structural data and support development of a model capable of predicting state failure. A summary of the strengths and weaknesses of the data streams is provided below in

Table 1.

Table 1: Summary Table Data Streams: Strengths and Weaknesses

Summary Table Data Streams: Strengths and Weaknesses		
Type of data	Strengths	Weaknesses
Structural data	-Comparable temporally and spatially -Generalizable beyond a specific country -Reliable -In wide use by aid agencies -Considered essential for desk officers for doing strategic analysis -Benchmarking possible -Lends itself to statistical analysis --The basis for long term forecasting and modelling	-Difficult to operationalize unless one understands the underlying concepts -Available data can be limited, particularly at sub-national levels -Desired disaggregation (e.g. demographic data) often not available -Data lag can be an issue -Identifies broad patterns but not necessarily context
Events	-Widely available information sources -Up to date information and real time analysis possible -Machine and human coding can be cross-checked -Lends itself to impact assessment and programme evaluation -Provides basis for developing a short to medium predictive and forecasting capacity -Provides basis for developing scenarios	-Context must be matched with stakeholder analysis -Potential bias of sources -On balance, machine-coded events are less resource-demanding than human-coding, but it lacks contextual basis for accurate assessment
Expert opinions	-Detailed in-depth country knowledge -Access to broad-based knowledge -Can highlight unseen but important concerns and opportunities -Systematic comparisons provide the basis for forecasting and risk analysis -Most useful as a cross check against other sources of information	-Expert Subjectivity -High costs -Development of questionnaires must be fitted to country issues, capacity and concerns -In-country coordination raises questions of consent and costs of implementation within host country

Assessment at a distance is perhaps a second best but has proven cost effective in comparison to training and deploying field monitors and, arguably, encourages objectivity. Start-up costs aside, machine coded events analysis is arguably the least costly, though one loses the ability to relate and incorporate context. In brief, human coding has the advantage of maintaining a global standard of decision rules while not precluding state- and situation-specific analysis.

It was appreciated from the onset that events-based analysis cannot provide a complete description and explanation of the context relating to complex situations, nor does it necessarily provide a complete

representation of root causes. Elucidation and interpretation must come from Subject Matter Experts (SMEs). Their forte and key contribution lies in helping to identify COIs and associated events that denote specific interactions among key actors (e.g., stakeholders) in a given situation.

3.3 Expert Opinion

Human insight can offer invaluably nuanced views of a country's stability. Qualitative information, of this kind is an indispensable complement to the systematic collection of statistical data, as it uncovers details and fine distinction. Put simply, when correctly structured, events data in conjunction with analysis from expert opinion can provide the "why" behind the "what" revealed through structural and dynamic data analysis. SMEs can provide detailed insight into specific issue areas, as well as offer ideas about rationale, motivation, causal linkages and what areas deserve the most attention going forward, either because they are functioning well and can be used to propagate positive reform in other parts of the governance system, or because they are weakening and threaten to undermine stability and development in other sectors.

The challenge lies in integrating structured and unstructured data – optimizing the benefits of holistic, numerical and thematic analyses. Narrative summation suffers from selective explanation and, as noted earlier, statistical analyses are limited by the ways that they can be and are operationalized. While qualitative information attempts to minimize this problem by describing the whole of a situation in detail, including all the bits and pieces that are difficult to include in a statistical analysis, by itself it is often insufficient. A human rights expert with long experience in a country can provide a full picture of the local rights environment, bringing in elements of culture, history, and analogous situations but ideally this would be combined with quantitative indicators. A thematic approach, using ordered scoring and exploiting inferential human scoring, seems to offer middle ground and best value.

Individual expert opinion tends to reveal only one part of a larger picture. An expert panel could include individuals that approach different elements of country performance from different ideological and professional perspectives. They could be used as a challenge and support function to the quantitative analysis embedded in this project. Still people have subjective viewpoints, whether they are ordinary citizens or specialists. Structuring the analysis is part of the solution but even the Political, Military, Economic, Social, Infrastructure, and Information (PMESII) framework currently in vogue comes with risk. PMESII specialists may overestimate the importance of their field to the overall situation in a country. Ideology can also cloud opinions, as can personal experience and bias. In brief, by and of itself a framework cannot provide an objectively true description of a country's performance and prognosis. A combination of people, process and tools are required, and all data streams must be exploited.[37]

3.4 Combining Data Streams

Meaningful analysis will include as much information as possible, ideally from all three of the sources described above. There are three main benefits to combining information streams:

[37] For more details on the attributes of each type, see Wyjad, K. 2007 CIFP Democracy and Governance Handbook (available at www.carleton.ca/cifp).

- Challenge function: If different streams tell different stories one can investigate further to discover which is correct; analysis based upon a single information stream is impossible to verify or validate.

- Depth of coverage: No single stream can provide a complete picture of a country's performance. Combining streams can give more information at different levels.

- Finer focus: Looking at different streams lets the analyst see more clearly what factors contribute most to outcomes. This idea connects to the previous two advantages of multi-stream analysis. If all three streams suggest that a certain factor is the key driver of a certain phenomenon one can have confidence that this is the case. However, if there is disagreement one can investigate further to gain clarity and find out which is correct. Likewise, deeper coverage will allow a more detailed picture of how various factors interact to produce outcomes, allowing causal relationships to go beyond simple cause and effect to embrace network dynamics.

3.5 Trend Analysis

The focus in examining structural factors is on discerning variability between states. The focus in examining events data is on discerning situational variability within a state. Country analysis and assessment of state fragility must consider both the current state as well as the direction the country is moving. Is the state on a downward trajectory and thus on the path to failure? Or is it entering a potentially destabilizing period of reform? A nation that is emerging from a civil war may exhibit extremely poor performance in almost all areas if examined as a snapshot frozen in time but, when dominant trends are considered, the country may be more appropriately deemed to be in the midst of a rapid recovery made possible by its newfound organization. Conversely, a democratic system may appear superficially strong and stable until one takes into account, for example, a growing and unchecked trend toward authoritarian executive behaviour. Timely analysis is required to support operational decisions. Trends are most easily and reliably visible through dynamic data analysis but, to a certain extent, tempering this with expert opinion is desirable if not required. From a programming perspective, trends are of crucial importance; understanding them allows decision makers to focus on areas showing worrisome deterioration, or to take full advantage of positive momentum as a resource multiplier to support a crucial governance sector.

As a final note, one must always understand and appreciate collaborative and conflicting processes as distinctive arrangements reflective of unique structures. Institutions and processes vary significantly between countries and societies. Assessing processes goes far beyond going through a checklist to identify whether or not a predetermined ideal model exists in a given country. Rather, assessing these processes involves identifying how (and if) a country's political system functions and interacts with its citizenry, and then determining what areas could stand improvement or reinforcement, and what areas are working effectively by satisfying the desires of the related population. Developing a model capable of measuring fragility and predicting state failure is tested given these factors. The approach and methodology adopted leveraged the prior work cited and attempts to address the related challenges.

4 Approach and Methodology

Two approaches were adopted in this study for quantitative early warning analysis. The first involved collecting annual data on structural factors, such as GDP growth. These variables provide knowledge of the root causes of instability, the foundations of intrastate conflict. However, structural factors cannot explain whether a structurally fragile country is at greater or lesser risk of instability for any given month, week, or day. The second approach provides more dynamic data for early warning, through events-based analysis. This method allows analysts to have a better sense of real-time changes, which gives a more precise indication of when a country is more prone to conflict. The combination of the two is required to support policy formulation. Essentially, structural data are important for understanding *which* states are at risk, and events data are key to knowing *when* a state is close to instability/failure.

Taken together, structural and events analysis can provide the comprehensive awareness necessary for proper early warning regarding state failure. CIFP has combined these methods qualitatively in order to provide "context-rich country assessments that are nonetheless still comparable against the performance of peers".[38] One area not yet addressed was the quantitative combination of structural and events-based data. A brief explanation of the associated conceptual EWM is given followed by a description of the CIFP approach to structural analysis and events-monitoring. The final sub-section describes the overall project methodology, which uses the CIFP approach and two first-generation approaches (PITF, VRA).

4.1 Conceptual Model

The conceptual EWM combines state structural factors and events data to forecast a crisis of interest (COI). Structural factors establish a baseline – and initial filter - that must be factored in and weighed against the occurrence and impact of incidents. Events data are incorporated to generate an aggregate Index of State Tension (IST) reflecting and linking long term and short term contributors to state fragility. A numerical value is derived for a state as a summary representation of the level of state tension. The structural component is not subject to rapid changes and forms a baseline condition. Although there remains variance in factors used to calculate structural fragility there is emergent consensus, correlation and convergence. GDP, infant mortality, trade imbalance and other similar variables are objective examples that comprise the structural factors. Events can be characterized as dynamic 'accelerators' or 'decelerators' that give rise or fall to state tension and, as described, events monitoring provides accessible, the time-sensitive data. Given that the latter can often vary in intensity, frequency, and nature (whether stabilizing or de-stabilizing), numerous possibilities exist to characterize the cumulative impact of events on state tension (escalation of or towards instability). Temporal considerations pose a complicating factor. For example, it is expected that recent events should weigh more heavily than past events, but to what extent is subject to statistical analysis. A state will pass from stability into a state of vulnerability, and later lapse into state failure, if the value assigned in the IST crosses pre-determined thresholds. It was hypothesized that these thresholds could be determined through regression analyses of historical data, and that patterns preceding a COI could be identified.

Logistic regression (LR) is a statistical technique that can be used to predict the probability of an event occurrence by fitting data to a logistic curve. Regression refers to the use of several predictor variables. LR is a convenient way of describing the relationship between independent variables and a dependent variable expressed as a binary response. Given that forecasts of a COI are probabilistic, LR is an appropriate statistical model to exploit the data collected.

[38] Carment, D; Gazo, J; Prest, S. *"Risk Assessment and State Failure"*. Global Society. Vol 21, No 1. Pp 47-69.

The model can be expressed mathematically i.e. logit is the LR function that comprises the independent variables:

$$\text{logit} = \gamma \cdot (\text{IST} - \alpha + \beta_2 \cdot E_{last})$$

where γ = shape parameter,
α = offset parameter,
β = weight parameter,
IST = Index of State Tension
$= SF + \beta_1 \cdot E_{past}$
SF = structural factors
E_{past} = history of past events
E_{last} = most recent events

The probability of a COI occurring is given by: $e^{logit}/(1 + e^{logit})$

The logit offers a useful discrimination between those states in imminent danger of failure/conflict and those that are not. The transition of the probability of a COI occurring from 0 to 1 is the zone of interest where the "tipping point" might be identified.

The logit function comprises scores for past events, E_{past}, and for the most recent events, E_{last}. Not prescribed, however, is the duration for the assessment of past events, which was left as an exploratory option and investigated in fitting the model to the data. Recent events were considered on a much shorter time scale in view of the anticipated application of the model. That is, it was envisioned that when real-time events are entered into the model, the duration of E_{last} should be sufficiently small so that it can be frequently refreshed (e.g., every 2 weeks). Thus, the model could be calibrated with events data both of the past (duration to be determined) and of the most recent period preceding a COI. All of these aspects of the model are illustrated, starting with
Figure ***12***.

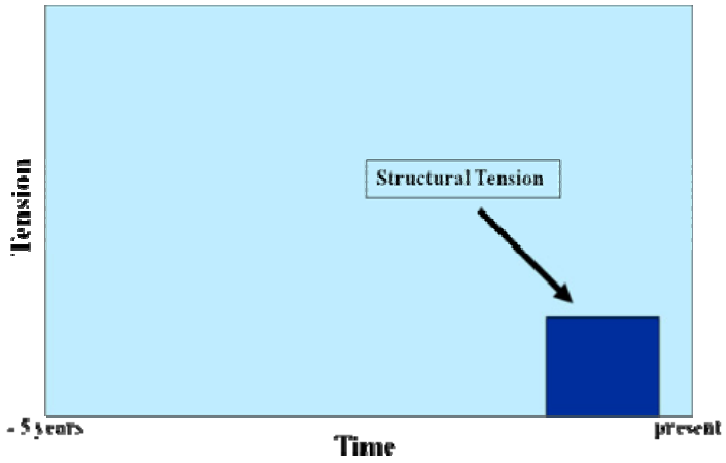

Figure 12: Conceptual Model

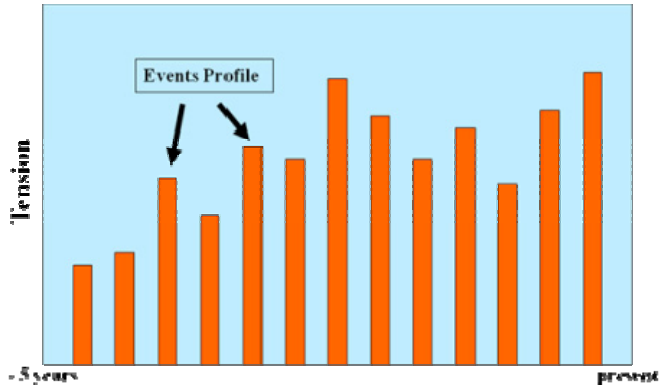

Figure 13: Conceptual Model

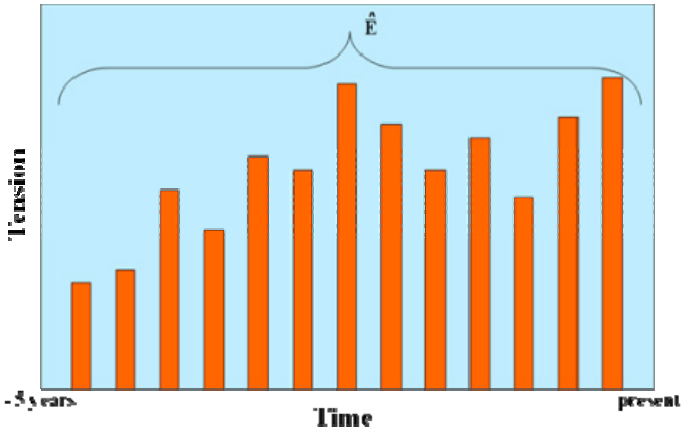

Figure 14: Conceptual Model (E is Epast)

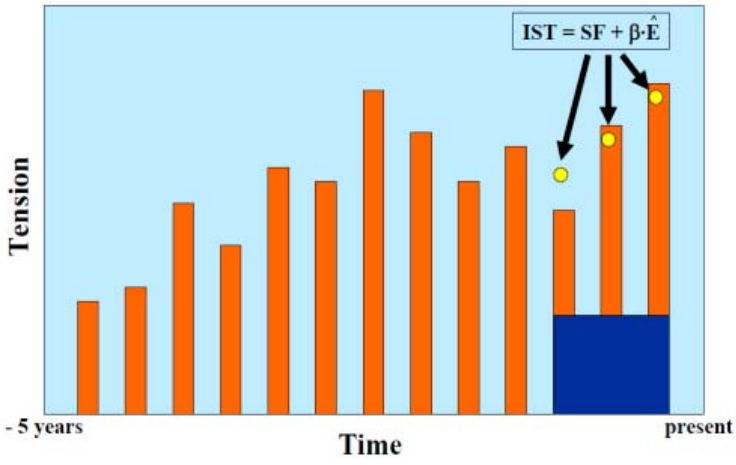

Figure 15: Conceptual Model (E is Epast)

The figure below (Figure 16) depicts the transition of a state through various degrees of 'tension'. The dark blue box represents the contribution of structural factors (i.e., based on political, economic, and

social indicators) to state tension. The orange bars represent the contribution of events, which can be stabilizing or de-stabilizing, to state tension. The yellow dots represent an aggregate i.e. the combination of these (weighted) contributions to state tension. When its value resides below the horizontal orange line, the state is considered either stable (i.e. highly unlikely to de-stabilize even in the event of violent episodes) or at 'alert' (i.e. violent events might elevate the state into the vulnerable zone). A state in the vulnerable zone can transition into instability (unstable zone) if events are sufficiently potent, in which case such events are labeled as 'triggers', or the cumulative effect is sufficient to create a COI. Events precipitate reactions and generate pre-conditions for crises. States in the vulnerable zone warrant early warning and were the focus of this study. As depicted below, a state above the red line is considered unstable and its likelihood of failing is quite high, which warrants an urgent warning (UW). Conceptually the orange and red line thresholds are defined by model parameters (Γ, γ, α, β) that can be estimated using logistic regression.

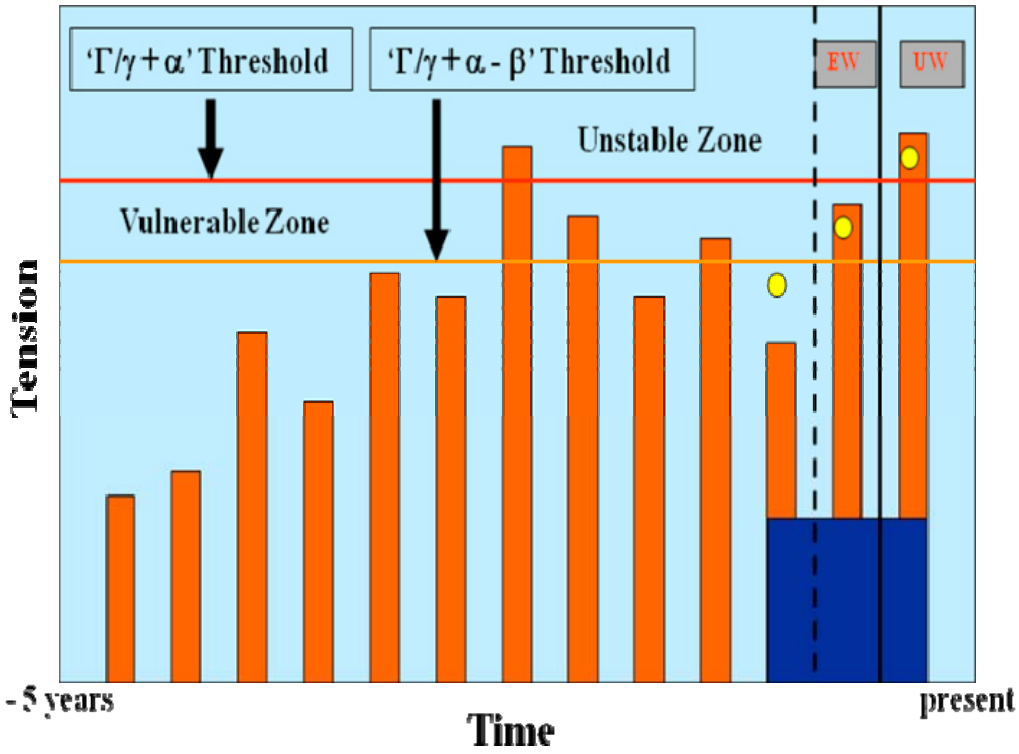

Figure 16: State Transition

4.2 Structural Data

All variables must be quantified to support statistical regression. For structural factors, this involved determining a metric that scores a country's status against several indicators of performance over time. As discussed earlier, the intent was to take a holistic view of state failure and consider a broad range of factors including governance, economy, social conditions, security, etc. The structural data are important because it provided an initial filter identifying the instances of state failure warranting detailed investigation i.e. it was used to select the suite of case studies. Two factor sets and associated metrics were considered; the logit function of state condition developed by the Political Instability Task Force (PITF) and the fragility index (FI) developed by CIFP (see www.carleton.ca/cifp).

PITF

The PITF has identified a set of seven variables used to predict state failure' (identified as adverse regime change, revolutionary war, genocide/politicide, and/or ethnic war) three years into the future. Regime consistency, infant mortality, economic openness, militarization, neighbourhood war, autocracy, and partial democracy are significant structural variables and are assessed annually. Parameter estimates (weights) of these variables are determined by PITF using LR and are used to generate a logit value for each country of interest using open-source data. In essence, the more positive the logit value, the more likely the state would experience failure. These data are available from PITF (www. globalpolicy.gmu.edu/pitf); the values provided a suitable metric of a state's structural status for use in the study.

CIFP's Authority-Legitimacy-Capacity (ALC)

First-generation research on the structural causes of state fragility, such as the PITF project, defined state failure as the inability of a state to respond to widespread violence, attributable usually to political conflict between groups. This research has focused largely on mass violence caused by institutional weakness and failure, driven by experiences since the end of the Cold War (Somalia, Rwanda, Bosnia and Herzegovina etc).

CIFP has adopted a broader approach and assumes that failure can occur for a number of different reasons. State instability is viewed as a deficiency in the authority, legitimacy and/or capacity[39] of a state to administer the provision of public goods and services. That is, the CIFP model contends that a state weakness and/or failure can be determined along three axes: authority, legitimacy and capacity (ALC).

> *Authority* refers to the ability of the state to enact binding legislation over its population and to provide the latter with a stable and safe environment. A state with weak authority is unable to provide a secure and stable environment and cannot enforce its laws. Typical proxies for authority include variables such as the level of corruption and contract regulation
>
> *Capacity* refers to the power of the state to mobilize public resources for productive uses.[40] Capacity indicators include measures such as GDP per capita, the trade balance and foreign aid receipts.
>
> *Legitimacy* refers to the ability of a state to generate public loyalty and support and acceptance of governance or specific policies. It is measured by variables such as regime type, human rights and gender empowerment.

[39] Gagne, Jean-Francois., Samy, Yiagadeesen., Carment, David., Prest, Stewart. *Small States, Resilience . and Governance*, July 2007. Presented at the May 2007 Small States and Governance Conference in Malta.
[40] Carment, Prest, Samy. 2009.

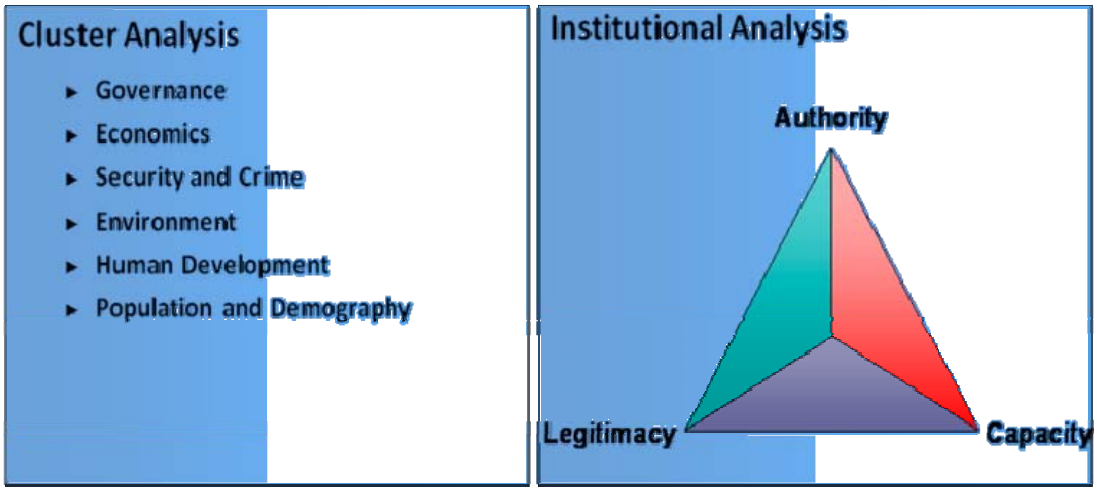

Figure 17: Cluster Analysis and Institutional Analysis[41]

Thus CIFP analysis of state fragility begins with the understanding that, to function effectively, a state must exhibit three fundamental properties: authority, legitimacy, and capacity. These terms are further explained in detail below, along with their implications for the analysis of state fragility and failure. These constructs reflect the functions of a state and its component parts.

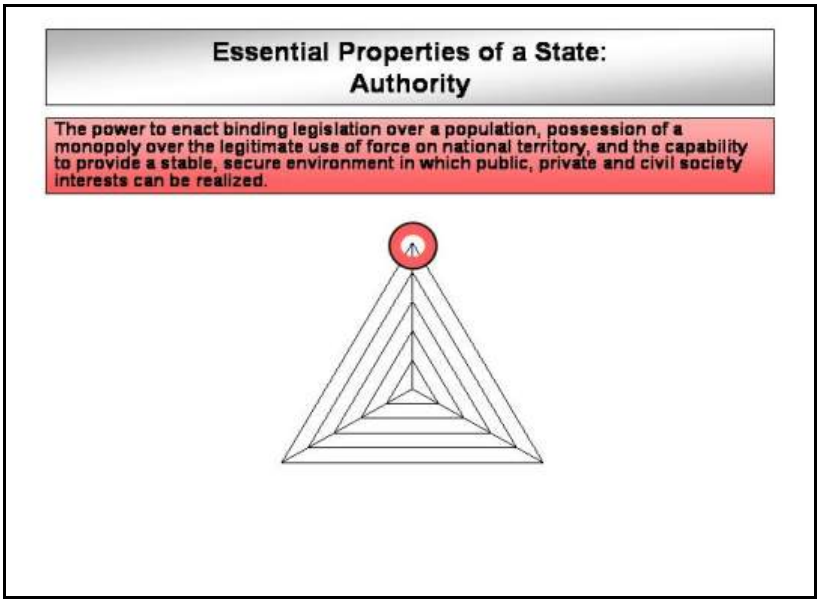

Figure 18: Essential Properties of a State – Authority

[41] Figures 17 through 21 are reproduced from the CIFP Concept Paper located at (www.carleton.ca/cifp)

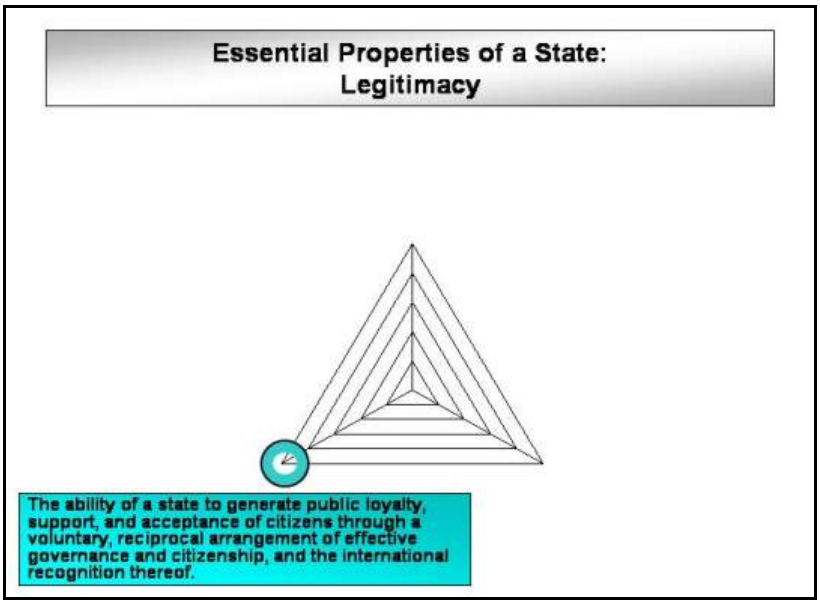

Figure 19: Essential Properties of a State - Legitimacy

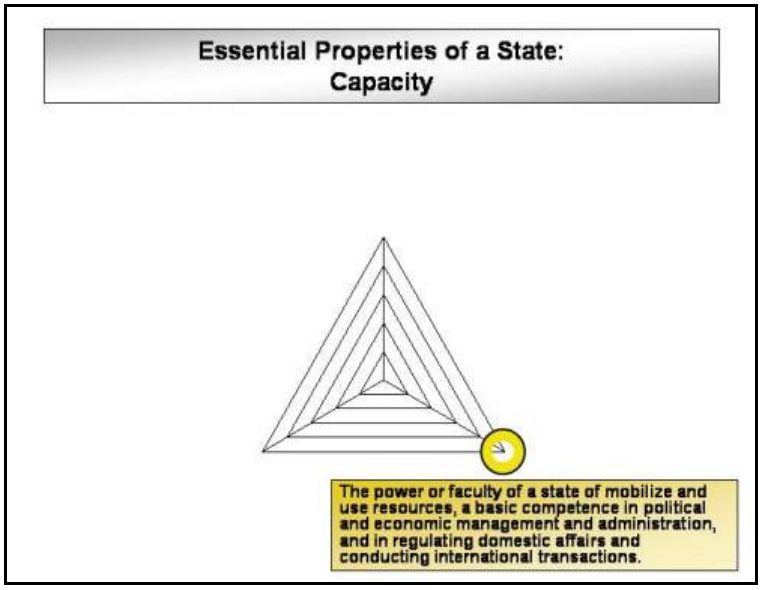

Figure 20: Essential Properties of a State – Capacity

The CIFP state fragility index (FI) uses relative rankings to list countries within each ALC category as well as an overall fragility ranking. In addition to the general ALC framework, CIFP also organizes data into six clusters that collectively characterize state performance along governance, economics, security and crime, environment, human development, and population and demography lines as shown in Figure 21.

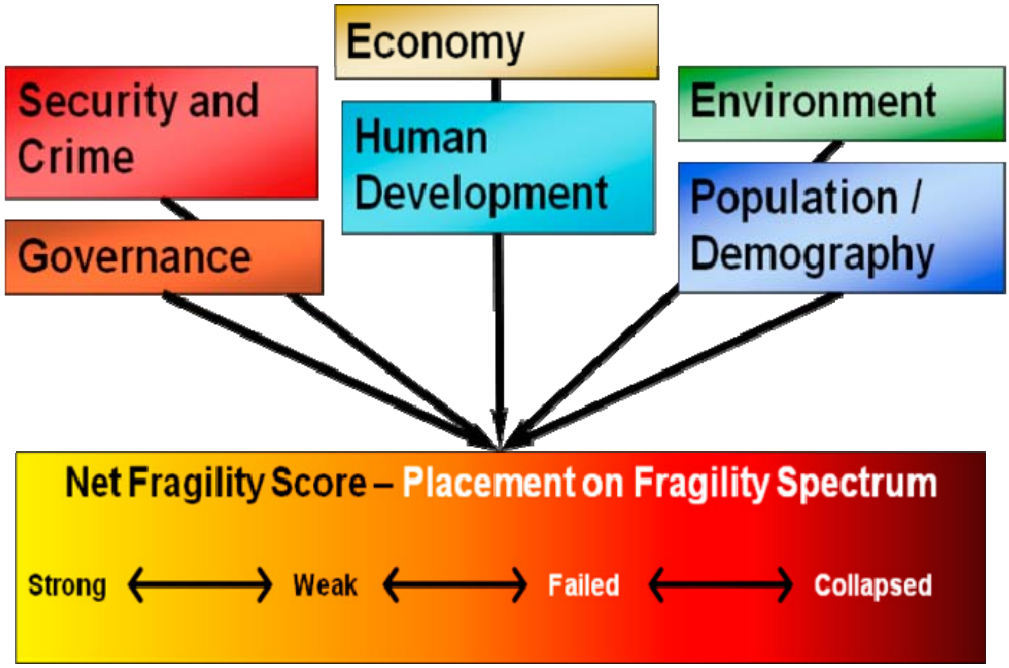

Figure 21: Net Fragility Score

Although the ALC framework creates a conceptual distinction and functional division between the three dimensions, in reality, the dimensions are not independent and a change in one is likely to be related to, or caused by, a change in another. Nevertheless, the discrimination can provide important insights. The importance of this multidimensional approach is evidenced by the case of North Korea, which had an overall ranking of 52 in 2007. However, on closer examination, North Korea emerged as extremely fragile in its legitimacy, which was balanced by better scores in authority and capacity. The ALC method thus identified North Korea as a fragile state due to frail legitimacy, whereas the state may not have been identified by other approaches that rely on mono-causal definitions of failure.

CIFP Structural Data

CIFP's state fragility index scoring table and Top 40 Fragile States are shown below in Table 2. An explanation describing derivation is in order. The analysis draws on data representing more than 70 indicators (culled from an initial listing of more than 100) selected on the basis of relation to state fragility and level of country coverage. In addition to the assessment of a state's relative levels of authority, legitimacy, and capacity, each country profile includes a cluster-based summary of state performance. Like its predecessor, the CIFP conflict risk index, the fragility index makes use of comparative assessment. In ranking state performance relative to a given indicator, global scores are divided into nine equal groups and converted to a 9-point index. The best performing 9th of states receive a score of 1, the second 9th a score of 2, and so on. For example, the countries with the highest GDP per capita score a 1, while those with the lowest GDP score a 9. Since relative country performance can vary significantly from year to year – as in the case of economic shocks, natural disasters, and other externalities – averages are taken for global rank scores over a five-year time frame.

The most recent five years contained in the CIFP data set are used in generating this index. Scores may be adjusted to capture positive or negative trend lines, and reflect excessive volatility as well. Once all scores have been recorded, the results for a given country are averaged to produce its final standing. As portrayed in

Table *2*, in general, a high score – 6.5 or higher – indicates that a country is performing poorly relative to other states. This approximates the upper quintile of countries above the global mean. Such a score may be indicative of an arbitrary and autocratic government, poor economic performance, low levels of human development, or the presence of a destabilizing structural condition such as a significant youth bulge or a critical lack of arable land. A low score – in the range of 1 to 3.5 – indicates that a country is performing well relative to others, or that a country's structural conditions present little cause for concern. Values in the moderate 3.5 to 6.5 range indicate performance approaching the global mean.

Table 2: Fragility Index Scoring Scale

Score	**Description**
1-3.5	**Country performing well relative to others**
3.5-6.5	**Country performing at or around the median**
6.5+	**Country performing poorly relative to others**

Table 3 lists the 40 top fragile states using data up to 2006 as determined via the CIFP Fragility Index. There are many different organizations who produce such annual ratings; Foreign Policy's annual Failed State Index for example. The CIFP Fragility Index in the table below is a representative example of such a ranking system. For each state, the table includes a net fragility score, ALC scores and cluster scores. Table 4 provides a list of the 20 most fragile states within CIFP authority, legitimacy, and capacity sets. These tables provide a clear illustration of the multifaceted nature of state fragility. While some states display weakness along virtually all dimensions, the situation for most is more complex, with states exhibiting unique elements of both stability and fragility. On the basis of the fragility index, one may not only identify broad areas of relative strength or weakness, but also derive detailed country profiles to pin point the precise source of the phenomenon. This "drill-down" capability in turn can be used to support programmatic decisions and to identify areas or trends of concern that require supplemental investigation and/or further monitoring. In addition, the information provides a framework with which to evaluate policy effectiveness.

Ultimately, the results in Tables 3, 4 and 5 provide an important validation of the framework, and illustrate its ability to capture the breadth and depth of state performance.

Table 3: The CIFP Index - Top 40 Fragile States for 2006[42]

[42] Notes these date tables are reproduced from Carment et. al. 2009.

Country	Fragility Index	ALC Scores			Indicator Clusters					
	F	A	L	C	Govern-ance	Economics	Security & Crime	Human Development	Demography	Environment
Sudan	7.02	7.35	6.88	6.75	7.86	6.62	8.03	7.85	6.51	5.27
Afghanistan	6.93	7.40	7.09	6.24	6.63	7.01	8.30	7.75	4.85	5.42
Congo, Dem. Rep.	6.91	7.12	6.56	6.76	7.96	6.56	6.81	8.42	7.50	4.00
Somalia	6.89	6.79	7.15	6.54	7.57	7.05	6.41	8.36	6.05	7.06
Ethiopia	6.79	6.84	6.26	6.89	6.74	6.03	7.79	8.11	7.31	5.83
Iraq	6.77	7.26	7.17	5.38	7.75	6.44	8.31	7.36	4.64	4.74
Burundi	6.69	7.20	6.19	6.54	6.92	6.11	7.86	8.43	6.89	4.07
Haiti	6.65	6.83	6.72	6.50	7.42	7.00	6.25	6.76	6.21	5.50
Pakistan	6.60	6.74	5.95	6.45	6.81	5.38	8.30	7.93	6.68	6.20
Liberia	6.59	6.70	6.55	6.62	7.34	7.02	5.96	7.41	6.86	4.45
Eritrea	6.57	6.41	6.87	6.46	6.90	7.13	5.92	7.95	6.54	4.35
West Bank and Gaza	6.57	6.59	6.68	6.62	7.46	7.34	6.53	4.40	6.26	7.20
Cote d'Ivoire	6.55	6.81	6.17	6.19	7.62	5.73	6.92	8.36	7.14	4.29
Nepal	6.55	7.02	6.50	6.20	7.11	6.36	7.59	6.23	6.59	5.46
Nigeria	6.53	6.82	6.06	6.21	6.83	5.80	6.55	7.63	7.58	5.58
Yemen, Rep.	6.45	6.20	6.11	6.29	7.18	5.64	6.20	7.77	6.95	5.78
Chad	6.44	6.35	6.40	6.50	7.43	6.39	5.48	8.45	6.56	3.82
Angola	6.28	6.88	5.84	5.67	7.04	5.55	6.76	8.57	6.96	4.35
Togo	6.27	5.90	6.62	6.33	7.19	6.81	4.41	7.10	6.44	4.83
Sierra Leone	6.22	6.01	6.13	6.71	6.62	6.82	4.05	8.75	7.11	4.30
Cen. Afr. Republic	6.21	6.09	6.36	6.17	7.30	6.46	4.91	8.65	6.57	2.24
Guinea-Bissau	6.19	6.26	6.25	6.27	6.58	6.80	4.55	8.09	7.07	3.82
Uganda	6.17	6.67	5.76	6.18	6.16	5.93	6.85	5.84	7.79	4.26
Myanmar (Burma)	6.15	6.82	6.45	5.30	7.33	6.20	7.29	5.67	5.13	4.70
Djibouti	6.12	5.10	6.12	6.77	6.69	6.27	4.12	7.89	5.88	5.79
Kenya	6.12	6.48	5.64	5.97	6.58	5.74	5.60	6.56	7.37	5.17
Zimbabwe	6.11	5.95	6.39	5.81	7.38	6.22	4.97	6.32	5.68	5.66
Comoros	6.08	4.87	6.10	6.98	6.25	7.01	2.82	7.64	6.56	5.04
Guinea	6.08	5.85	5.90	6.25	7.11	6.02	4.57	7.80	6.69	3.79
Equatorial Guinea	6.03	5.05	6.46	6.03	6.97	6.17	4.11	6.81	6.30	4.91
Bangladesh	6.02	6.28	5.52	6.07	6.46	5.76	6.31	6.16	6.40	5.24
Korea, North	6.02	5.28	6.69	5.78	7.27	6.63	4.16	6.82	4.85	6.52
Tajikistan	5.98	5.30	6.23	6.24	6.92	6.24	4.86	5.91	5.43	6.02
Cameroon	5.96	5.85	5.92	5.81	7.21	5.81	4.50	7.03	6.87	4.10
Iran	5.95	6.42	6.09	4.88	7.41	5.38	6.95	4.65	5.41	6.03
Niger	5.93	5.26	5.98	6.59	5.94	6.47	3.26	8.77	6.73	3.67
Congo, Rep.	5.92	5.83	5.81	5.74	7.13	5.55	4.71	7.42	6.34	4.39
Rwanda	5.91	5.95	5.75	6.03	6.56	6.06	5.25	6.35	7.06	3.35
Mauritania	5.90	5.08	5.86	6.38	6.03	5.82	5.13	7.07	5.41	6.13
Laos	5.87	5.94	6.05	5.53	6.33	6.04	5.50	6.62	6.16	3.89

Table 4: Twenty Most Fragile States by ALC Component for 2006

Authority		Legitimacy		Capacity	
Afghanistan	7.40	Iraq	7.17	Comoros	6.98
Sudan	7.35	Somalia	7.15	Ethiopia	6.89
Iraq	7.26	Afghanistan	7.09	Djibouti	6.77
Burundi	7.20	Sudan	6.88	Congo, Dem. Rep.	6.76
Congo, Dem. Rep.	7.12	Eritrea	6.87	Sudan	6.75
Nepal	7.02	Haiti	6.72	Sierra Leone	6.71
Angola	6.88	Korea, North	6.69	Liberia	6.62
Ethiopia	6.84	West Bank and Gaza	6.68	West Bank and Gaza	6.62
Haiti	6.83	Togo	6.62	Niger	6.59
Myanmar (Burma)	6.82	Congo, Dem. Rep.	6.56	Benin	6.55
Nigeria	6.82	Liberia	6.55	Burundi	6.54
Cote d'Ivoire	6.81	Nepal	6.50	Somalia	6.54
Somalia	6.79	Equatorial Guinea	6.46	Chad	6.50
Pakistan	6.74	Myanmar (Burma)	6.45	Haiti	6.50
Liberia	6.70	Chad	6.40	Eritrea	6.46
Uganda	6.67	Zimbabwe	6.39	Mozambique	6.46
West Bank and Gaza	6.59	Central African Republic	6.36	Pakistan	6.45
Kenya	6.48	Ethiopia	6.26	Mali	6.42
India	6.45	Guinea-Bissau	6.25	Burkina Faso	6.40
Iran	6.42	Tajikistan	6.23	Mauritania	6.38

Table 5: Twenty Most Fragile States by Cluster for 2006

Governance		Economics		Security&Crime	
Congo, Dem. Rep.	7.96	Timor-Leste	7.92	Iraq	8.31
Sudan	7.86	West Bank and Gaza	7.34	Afghanistan	8.30
Iraq	7.75	Eritrea	7.13	Pakistan	8.30
Uzbekistan	7.67	Somalia	7.05	Russia	8.18
Cote d'Ivoire	7.62	Liberia	7.02	Sudan	8.03
Somalia	7.57	Afghanistan	7.01	India	7.94
West Bank and Gaza	7.46	Comoros	7.01	Burundi	7.86
Turkmenistan	7.44	Haiti	7.00	Ethiopia	7.79
Chad	7.43	Tonga	6.89	Colombia	7.68
Haiti	7.42	Kiribati	6.86	Nepal	7.59
Iran	7.41	Sierra Leone	6.82	Turkey	7.50
Zimbabwe	7.38	Togo	6.81	Israel	7.42
Liberia	7.34	Guinea-Bissau	6.80	Indonesia	7.32
Myanmar (Burma)	7.33	Sao Tome and Principe	6.77	Myanmar (Burma)	7.29
Central African Republic	7.30	Nicaragua	6.76	Algeria	7.23
Korea, North	7.27	Solomon Islands	6.66	Sri Lanka	7.22
Libya	7.23	Korea, North	6.63	Iran	6.95
Cameroon	7.21	Gambia	6.62	Cote d'Ivoire	6.92
Togo	7.19	Sudan	6.62	Uganda	6.85
Yemen, Rep.	7.18	Congo, Dem. Rep.	6.56	China	6.83

Human Development		Demography		Environment	
Niger	8.77	Uganda	7.79	Singapore	7.56
Sierra Leone	8.75	Malawi	7.67	New Caledonia	7.47
Central African Republic	8.65	Nigeria	7.58	Oman	7.39
Angola	8.57	Congo, Dem. Rep.	7.5	Bahrain	7.35
Chad	8.45	Kenya	7.37	West Bank and Gaza	7.20
Burundi	8.43	Ethiopia	7.31	Somalia	7.06
Congo, Dem. Rep.	8.42	Cote d'Ivoire	7.14	Malta	7.00

Cote d'Ivoire	8.36	Burkina Faso	7.13	Qatar	6.99
Somaila	8.36	Sierra Leone	7.11	United Arab Emirates	6.93
Burkina Faso	8.29	Guinea-Bissau	7.07	Bahamas	6.88
Mali	8.19	Rwanda	7.06	Saint Lucia	6.68
Ethiopia	8.11	Angola	6.96	Seychelles	6.62
Guinea-Bissau	8.09	Yemen, Rep.	6.95	Kuwait	6.56
Mozambique	8.04	Burundi	6.89	Korea, North	6.52
Eritrea	7.95	Cameroon	6.87	Antigua and Barbuda	6.50
Pakistan	7.93	Liberia	6.86	Isle of Man	6.50
Benin	7.91	Guatemala	6.8	Puerto Rico	6.49
Djibouti	7.89	Niger	6.73	Korea, South	6.31
Sudan	7.85	Mali	6.7	South Africa	6.29
Guinea	7.8	Guinea	6.69	French Polynesia	6.25

These tables reveal both the diversity of the situations and challenges fragile states face and the disadvantage in grouping failed states using a "one size fits all approach which is common other datasets. To illustrate, a number of sub-Saharan African nations face serious problems arising from limited capacity. Some countries – including Mozambique, Mali, and several others that perform relatively well in areas of authority and legitimacy – face enormous challenges in terms of state capacity. Others, such as Colombia and Sri Lanka, score poorly in the area of authority as a result of the security challenges they face, but perform relatively well in measures of capacity and legitimacy. Still others, such as Belarus, face challenges to government legitimacy even as they continue to maintain some degree of state authority and capacity.

Certain countries face challenges in specific cluster areas. For instance, despite the presence of reliable governing institutions and robust economic development, many small states, particularly island nations such as St. Lucia, exhibit high levels of environmental stress. While such countries may benefit from assistance provided by members of the international community, the nature of that assistance and its method of delivery clearly will vary widely in each case. While a number of states appear on more than one list, only three – the DRC, Eritrea, and Haiti – appear on all three. This intriguing finding underscores the range of ways in which states exhibit fragility, and exposes the advantages of the ALC methodology in isolating and clarifying those varied experiences. These three states face challenges beyond those of other developing states; should members of the international community hope to engage such states effectively, their efforts must be carefully planned, appropriately coordinated, and well funded.

Research and experience suggests that states that appear on one or two of the lists also face unique challenges that require carefully tailored policy approaches. International engagement in Iraq, Colombia, or Sri Lanka obviously must take careful note of the volatile security situations in each country. Conversely, international development programs in states such as Mali, Mozambique, and Chad must make government capacity and human development a priority, aside from any considerations of security. Yemen is among the top 20 fragile states in terms of its legitimacy gap. It is also among the top 20 fragile states in terms of governance, demography, and environment. Clearly, any development program that does not take into account all these areas of state weakness has little hope of success. This is consistent with Canada's hallmark declaration of taking a comprehensive response to stabilizing fragile/failed states. [43] Simply put, policymakers must take the particular pattern of fragility in a given state into account when

[43] Department of Foreign Affairs and International Trade. *Canada's International Policy Statement: A Role of Pride and Influence in the World.* Her Majesty the Queen in Right of Canada: 2005.

crafting their engagement strategies; to do otherwise is to invite ineffective and perhaps destabilizing policy.

More importantly in the context of this project, the diversity between and within lists is notable, reinforcing the point that no single index can capture the full measurement of state performance i.e. present a complete picture. No country appears on all six lists, or even on five. While a number of countries appear on two or three lists, only a few – including Ethiopia, Somalia, and Burundi – appear on four different lists. The difference across the lists belies what each is designed to measure. A country that experiences significant poverty may appear on a list with a series of other countries that is attempting to measure poor aggregate economic performance, but will not necessarily qualify as a failed or fragile state if no violence or systemic government weakness exists. Conversely, a nation with comparatively greater economic prosperity may qualify as a failed state because of systemic, targeted violence against a particular diaspora or identity group. Each list will only show countries that satisfy the metrics being used, not a total list of the most failed and fragile states. When considering different perspectives on state failure and fragility, how then to identify an appropriate set of case studies – one which would reflect this diversity? Table 6 and Table 7 below represent the ALC measures and their subcomponents, and comparing a single fragility index figure with the ALC measures respectively.

Table 6: Correlation of Fragility Index Components

Correlation of fragility index components					
Index component	**No. of indicators**	**Fragility Index**	**Authority**	**Legitimacy**	**Capacity**
Authority	24	0.83			
Legitimacy	16	0.84	0.58		
Capacity	34	0.94	0.62	0.75	
Gender	8	0.74	0.40	0.75	0.77
Governance	12	0.77	0.57	0.93	0.64
Economics	19	0.79	0.49	0.75	0.85
Security and Crime	13	0.68	0.91	0.56	0.41
Human Development	19	0.91	0.69	0.63	0.94
Demographics	8	0.83	0.60	0.61	0.87
Environment	3	0.29	0.11	0.31	0.33

Table 7: Correlation of Selected Indicators to Fragility Index

Country	Fragility Index	Authority	Legitimacy	Capacity
Absolute poverty (% of population living on less than $1 a day, World Bank, data taken from most recent year)	0.66	0.36	0.33	0.78
CIRI empowerment human rights index (2004)	-0.54	-0.48	-0.69	-0.38
CIRI physical integrity human rights index (2004)	-0.60	-0.73	-0.54	-0.41
Freedom House Press Freedom Index (2004)	0.65	0.55	0.81	0.50
Freedom House civil and political rights index (2003)	0.64	0.52	0.78	0.51
Fund for Peace failed state index (2006)	0.89	0.79	0.89	0.78
Gender development index (UNDP, 2003)	-0.91	-0.68	-0.75	-0.96
GDP per capita (WDI, 2003)	-0.85	-0.60	-0.86	-0.82
Gender empowerment measure (UNDP, 2003)	-0.82	-0.67	-0.86	-0.75
GINI (WDI, most recent year)	0.45	0.31	0.35	0.49
Human development index (UNDP, 2003)	-0.90	-0.63	-0.70	-0.95
Infant mortality (2003)	0.81	0.58	0.62	0.85
Military spending (% GDP, 2002)	0.24	0.24	0.30	0.17
Foreign aid per capita (OECD, 2004)	-0.08	-0.25	-0.09	0.07
Foreign aid as % GNI (OECD, 2004)	0.48	0.17	0.29	0.60
Political stability (WB Governance database, 2004)	-0.81	-0.83	-0.75	-0.64
Polity IV Democracy-Autocracy index (2003)	-0.47	-0.27	-0.67	-0.40
Square of Polity Democracy-Autocracy index (2003)	-0.72	-0.53	-0.70	-0.70
SIPRI armed conflict database	0.21	0.38	0.08	0.12
Slum Population (% total urban)	0.78	0.61	0.56	0.80
Trade openness (2002)	-0.29	-0.34	-0.16	-0.27
Youth Unemployment (2001)	0.30	0.28	0.41	0.19

Note: Tables 6 and 7 are reproduced from Carment, El-Achkar, Prest and Samy (2006).

4.3 Case Selection

The importance of selecting an appropriate set of cases was appreciated from the onset. There were a number of inherent constraints, the most important being the availability of data, in particular 2 consecutive years of events data for both unstable and stable periods when fragility scores improved or deteriorated significantly. Events data were sought that compare computer and human coding. The cases chosen, and associated periods during which the highest changes in fragility scores were recorded over a one year period, are listed in Table 8,
Table *9* and Table 10).

The selection criteria initially proposed included identifying countries of interest based on consistently high fragility scores, volatility (fluctuations around the median), weakness (high score) in one ALC

dimension, regional distribution and Canadian interests (policy relevance). Next "country years" were determined based on the change in fragility in absolute terms from the previous year. The selection attempted to take into consideration several factors and involved reviewing countries with consistently high fragility scores, countries with volatile fragile scores, countries whose scores along one of the ALC dimensions were high, ensuring regional coverage, and considering where the Canadian government has historic ties or economic interests. Implementing the idea in practice proved problematic: 41% of the "stable" years fell 1 year before the "unstable" year, 10% of the "stable" years fell 1 year after the "unstable" year and 55% of the "stable years fell within 2 years of the "unstable" year. The selection criteria were revised to incorporate a requirement for a minimum of 2 years separation between "stable" and "unstable" years. A determination of data availability from PITF and VRA followed in order to assess the potential (in the long term) for conducting sensitivity analysis using different types of data. Table 8,
Table *9* and Table 10 list the final-selected cases. The numbers after each year refer to the FI for the year preceding the chosen year, and the year itself, respectively.

As inferred an iterative selection process was adopted to direct data collection. A number of cases had very obvious drivers contributing to failure, and stood out immediately. Elevated fragility scores and variation from the median tended to be the determinant factors.

Table 8: Primary Cases

COUNTRY	REASON FOR INCLUSION	UNSTABLE YEAR	STABLE YEAR
Afghanistan	High FI; High variation from median	1998 (6.90; 7.73)	2002 (7.73; 7.00)
Congo, Dem. Rep.	High FI; High variation from median	1996 (6.21; 6.65)	2002(7.01; 6.77)
Haiti	High FI; High variation from median	1994 (5.58; 5.98)	1998 (6.26; 6.06)
Somalia	High FI; High variation from median	1992 (5.61; 6.15)	1997 (6.60; 6.29)
Sudan	High FI; High variation from median	1994 (5.88; 6.55)	1997 (6.62; 6.53)
Pakistan	Different ALC	1994 (5.51; 5.94)	2003 (6.54; 6.35)
Sri Lanka	Different ALC	1994 (5.23; 5.61)	1997 (5.62; 5.28)

Table 9: Expanded (Secondary) List of Cases

COUNTRY	REASON FOR INCLUSION	UNSTABLE YEAR	STABLE YEAR
Albania	Variation around the median	1996 (4.80; 5.15)	1999 (5.15; 4.65)
Bahrain	Variation around the median	2006 (4.33; 4.71)	1995 (5.17; 4.76)
Bolivia	Variation around the median	1997 (4.77; 5.00)	1993 (5.10; 4.80)
Brazil	Variation around the median	1995 (4.60; 4.86)	2006 (4.51; 4.76)
Dominican Republic	Variation around the median; SIDS	1994 (4.71; 5.10)	1991 (5.40; 5.10)
El Salvador	Variation around the median	1996 (5.00; 5.18)	1992 (5.66; 5.31)
Gabon	Variation around the median	1992 (4.39; 5.03)	1995 (5.01; 4.72)

COUNTRY	REASON FOR INCLUSION	UNSTABLE YEAR	STABLE YEAR
Guyana	Variation around the median	1995 (4.43; 4.83)	1991 (5.15; 4.44)
Kazakhstan	Variation around the median	1998 (4.88; 5.24)	1992 (4.59; 4.41)
Libya	Variation around the median	1996 (4.55; 5.50)	1999 (5.69; 5.28)
Malaysia	Variation around the median; Regional	2001 (4.47; 4.66)	1996 (4.53; 4.36)
Mexico	Variation around the median	1994 (4.80; 5.37)	1997 (5.20; 4.84)
Moldova	Variation around the median	1992 (4.50; 4.92)	1995 (4.78; 4.59)
Paraguay	Variation around the median	2000 (4.74; 5.02)	2004 (4.95; 4.72)
Peru	Variation around the median	1992 (5.30; 5.71)	1997 (5.56; 5.30)
Qatar	Variation around the median	2001 (4.56; 4.97)	2004 (4.96;4.66)
Saudi Arabia	Variation around the median	2004 (4.81; 5.01)	1996 (5.17; 4.95)
South Africa	Variation around the median	1992 (5.18; 5.56)	1998 (5.08; 4.53)
Thailand	Variation around the median	1994 (4.50; 4.79)	2002 (4.50; 4.37)
Ukraine	Variation around the median	1994 (4.49; 4.83)	2005 (4.97; 4.69)
Venezuela	Variation around the median	1996 (4.79; 5.07)	2006 (4.99; 4.76)

Table 10: Expanded (Tertiary) List of Cases

COUNTRY	REASON FOR INCLUSION	UNSTABLE YEAR	STABLE YEAR
Chad	Escalation; suggested by one reviewer	2005 (6.35;7.08)	1993 (6.05; 5.83)
Cote d'Ivoire	Escalation; suggested by one reviewer	1999 (5.77; 6.09)	1996 (6.09; 5.66)
Ethiopia	Escalation; suggested by one reviewer	1994 (5.21; 6.24)	1998 (6.42; 6.26)
Jamaica	SIDS; suggested by one reviewer	2001 (4.53; 4.73)	2006 (4.81; 4.49)
Nigeria	Fragile: suggested by both reviewers	2002 (6.18; 6.48)	1999 (6.38; 6.20)
Papua New Guinea	SIDS; suggested by one reviewer	1998 (5.02; 5.48)	1991 (5.81; 5.23)
Yemen	Regional; suggested by one reviewer	1999 (6.10; 6.36)	1993 (6.30; 5.96)
Israel	Some variation; strategic/regional	1994 (4.57; 4.81)	1998 (4.65; 4.27)
Jordan	Variation around the median	1997 (4.77; 5.20)	1993 (5.03; 4.73)
West Bank and Gaza	Strategic/regional	2001 (5.68; 6.14)	1998 (6.54; 6.02)

Note: 1) F = Fragility; 2) SIDS = Small Island Developing States; 3) The numbers in brackets after each year refer to the FI for the year preceding that year and the year itself, respectively.

"Unstable" and "Stable" years in the above table refer to periods during which largest deteriorations or improvements (in an absolute sense) in fragility scores were recorded. In other words, even though an extremely fragile country would be considered unstable over several years, we still identified a particular period as "stable" if a large improvement was recorded in its fragility score. There are periods in which a state that qualifies as being 'failed' may equally be experiencing stability in a relative sense: more stable than usual, while still within the 'critical' threshold. This distinction between absolute and relative stability is noteworthy. Using these criteria it is possible for an extremely weak state to experience

periods of comparative stability. A conscious decision was made to examine the absolute change in scores, rather than using percentage changes.

Case Additions

As this phase of the project was concluding, it was determined that additional cases of state instability would be useful to further development of the EWM. The cases were drawn from the PITF and, hence, characterized as incidences of state failure owing to regime change, revolutionary war, ethnic war or a combination of these. These cases are listed below in Table 11.
Table 1 Again a randomly selected minimum of 20 events per month were coded and catalogued adding to the database.

Table 11 Additional Cases

Country	Period	Country	Period
Algeria	05/90 – 05/91	Kenya	10/90 – 10/91
Armenia	07/94 – 07/95	Lesotho	05/97 – 05/98
Belarus	04/94 – 04/95	Mali	06/89 – 06/90
Cambodia	07/96 – 07/97	Nepal	02/95 – 02/96
Central African Republic	03/02 – 03/03	Niger	01/95 – 01/96
Comoros	09/94 – 09/95	Rwanda	10/89 – 10/90
Congo-Brazzaville	04/96 – 06/97	Senegal	09/91 – 09/92
Congo-Kinshasa (DRC)	03/91 – 03/92	Sierra Leone	03/90 – 03/91
Egypt	02/91 – 02/92	Solomon Isles	06/99 – 06/00
Ethiopia (2 months)	01/99 – 02/99	Thailand (1 month)	07/03
(The) Gambia	07/93 – 07/94	Yemen	04/93 – 04/94
Haiti	09/90 – 09/91	Yugoslavia	02/97 – 02/98
Ivory Coast (6 months)	09/01 – 11/01 07/02 – 09/02	Zambia	11/95 – 11/96

A workshop on Formal and Behavioural Science Approaches to Understanding and Predicting Adversarial Intent was hosted by DRDC Toronto 16 February. In one of the presentations, Professor Peter Suedfeld (University of British Columbia) described research into integrative complexity i.e. using thematic content analysis, and scoring integrative complexity to forecast when leaders/elites have reached a decision point. One of his three case studies involved domestic violence in Zimbabwe in 2007-2008. It was recognized that this might prove an opportunity to link the two research thrusts and Zimbabwe was added to the Indicators of State Failure case list.

This additional research was commissioned under a separate contract but the information collected was added to the then existing database. Mention is included in this report to provide a complete documentary record. While this will support the continuing research effort, it must be underscored that these data were not available and not used in the initial analysis the results of which are reported below.

VRA- IDEA Events Data

The options available for obtaining events data included computer-aided extraction using an algorithm such as IDEA human-coding using historical records and SMEs. The former draws on downloadable data,

which were obtained from VRA for 1990-2005 inclusive.[44] It is noteworthy that this reduced the potential number of ISC (Intra-state conflict) cases preceded by at least 5 episode-free (as defined by PITF) years to 38 (e.g. 8 ethnic wars, 8 revolutionary wars, 14 regime changes, and 8 complex). Excluding events prior to 1990 drops the number of ISC cases from 38 to 21. If the period of required events data is relaxed to 2 years instead of 5, then the number of ISC cases increases to 30. A preliminary first-order analysis of the VRA events data for states identified by PITF that underwent ISC between 1990 and 2005 was conducted using three different event indices:

i) Instability (based on government and civil direct and forceful actions);
ii) Hostility (based on the WEIS (World Events Interaction Survey) events of demonstrations, sanctions, expulsions, seizure, and force), and
iii) Goldstein (1992; based on a conflict-cooperation scale).[45]

However, none of these approaches revealed any obvious predictive patterns leading to ISC.

CIFP Events Data

The focus of this project thus shifted towards developing a methodological framework for conducting SME-based events data analysis using the CIFP methodology.

Event Causality, Centrality and Intensity

To date, the focus of most events data coding has been on categorizing an event as cooperative or conflictual, not its contextualized impact, which requires interpretation and human judgement. Essentially, the research conducted under DRDC Toronto's auspices offers an alternative point of departure. It presented an effort to assess the accumulated impact of individual events while maintaining a standardized scale so as to facilitate comparisons of event trends over time, and even across cases. The objective of this program was to observe and report on events within a fragile state to provide a richer appreciation of the dynamics of fragility. This involved integrating structural, referential data providing a baseline with events data to provide a more comprehensive understanding.

An event is first identified from a credible source, such as Reuters and/or other on-line news sources, including local language and domestic press sources. For standardization and verification purposes, the event must be reported in a reputable and accessible source. It must also be situated in time and space, and will be coded according to an exact date and an exact location. Second, the event coding reflects interpretation i.e. it is catalogued by a human, not a machine. There is obvious merit in complementing news wire sources that are available through Lexis-Nexis, Google News and AlertNet etc. with trained assessment. As noted, the advantages of human coding, cited by those who use it, are that it provides for some contextual nuance in the process of evaluation and reduces the possibility of coding an event twice.

Following the CIFP methodology, once identified, events were graded using a 1 to 3 scale along three axes capturing causality, centrality and escalation dimensions. Itemisation and description of the factors is shown in ***Table 12***, which serves as a reference point for the remainder of the section.

Table 12: Dimensions of Composite Events Score

Score	Causal (Ca) Relevance

[44] Virtual Research Associates – VRA Reporter. http://vranet.com/productsRep.html
[45] Goldstein, J. *"A Conflict-Cooperation Scale for International Events Data."* Journal of Conflict Resolution. Vol 36. Pp 369-385

1	Event is relevant, but with no clearly delineable causal linkage to state stability or fragility (e.g. a funding announcement or an international soccer friendly).
2	Event is relevant, with a delineable, though indirect causal linkage to state stability or fragility (e.g. New legislation enhancing minority rights is passed, or a bomb detonates within an ethnically divided region).
3	Event is relevant with delineable and direct causal linkage to state stability or fragility. (e.g. Declaration of a ceasefire or assassination of a government minister.)
Score	**Centrality (Ce)**
1	Event affects less than 25% of political stakeholders.
2	Event affects 25% - 75% of political stakeholders.
3	Event affects more than 75% of political stakeholders.
Score	**Escalation (Es)**
1	Event is comparable to others experienced in the state in the previous six months.
2	Event is more intense than others experienced in the state in the previous six months.
3	Event is more intense than others experienced in the state in the previous five years.

Regarding causality, certain events may be significant but not conceptually linked to stabilization/destabilization in a direct and clearly delineable fashion. These events, such as the introduction of a new aid project, a soccer match between rival factions, or the expulsion of a small international NGO, are assigned causality ratings of 1. Alternatively, events are scored 2 when they possess a delineable link but only indirectly. For instance, the commencement of peace talks or a speech by a popular local politician inciting ethnic hatred exhibit a delineable but indirect link to stability: peace talks reduce risk due to confidence building or the production of a peace agreement and hate speeches can generate violent behaviour but not necessarily directly. Finally, when an event is clearly and directly related to the risk of instability such as a ceasefire, or the assassination of a government minister it would be assigned a causality measurement of 3.

The second component of an event is its escalation potential (Es). Es recognizes that events occur within the context of immediate history and that the escalation of events can affect the risk of future instability.

Regarding centrality, the analyst compiles a list of stakeholders, those individuals or groups that possess a discernible and politically oriented agenda, and are affected by events within a country. They often have an organizational structure and sufficient resources to affect state fragility. They may possess tangible resources and have the ability to affect fragility directly: for instance, the national leadership of the country. A stakeholder may also be a group of similar individuals without collective or concrete resources, but can affect events intangibly: for instance, ethnic or religious groups. Where there are multiple small stakeholders with broadly similar agendas, such as domestic humanitarian NGOs or rebel militias, such groups may be aggregated – however approximately – in similar collectives. For instance, though they are not identical, CARE and MSF possess sufficiently similar goals and behavioural characteristics that they may be grouped together as "international NGOs." Alternatively, where distinctions such as ethnicity or religion do not exist between civilians, there may be a broad "civilian/community" stakeholder.

Scoring on the three dimensions allows country contexts to be taken into account, and acknowledges that similar events can have very different magnitudes given regional realities. For instance, a ceasefire

arrangement may have a lower causality rating in a country that has experienced multiple failed ceasefires compared to a region that is experiencing its first ceasefire arrangement. Alternatively, assigning a centrality measurement to the disarmament of one paramilitary group depends on whether that group is the only paramilitary organization or whether there are hundreds of such groups. Finally, a car bombing would not be considered an escalation of risk in a region that experiences such bombings on a monthly or even weekly basis. However, a similar bombing would constitute a significant escalation of events if it occurred in a region that previously only experienced non-violent protest.

The third point to underscore is that an event cannot be an abstract or subjective process. For example, "the Pakistan government pressures terrorist groups to hand in weapons," is not an event but an abstract process. However, if the source of information goes on to specify observable steps in that process, for example "Pakistani government offers financial reimbursement for handed over weapons," then it is an event. Fourth, scaling for each event is assigned with due regard to local context. Whereas the FAST - VRA method, for example, applies a score based on how cooperative (0-10) or conflictual (-10 - 0) the event is, the CIFP approach assigns a stabilizing/destabilizing sign. A positive or negative score will be applied to each dimension of the composite events score such that all indicators are assigned a value according to their effect as stabilizing a state or destabilizing it respectively. The results can then be plotted against time to produce a general trend line. The composite indicator (Ca+Ce+Es) becomes the event score and can be used to create time-series trend lines, as event data are plotted over a given period of time. These trends can then be analysed in aggregate (data from all events are assessed) as well as disaggregated (events data are analysed by cluster) in an effort to understand the current trajectory of the state – whether it is increasingly stable, stagnating, or deteriorating into weakness and failure.

Hence each event detected and recorded by CIFP has a composite indicator statistic, which measures each event's magnitude, or more precisely, its net impact on state stability or fragility. To recap, the event score is calculated based on four contextual dimensions of an event, namely: whether the event is stabilizing or destabilizing; the relationship of the event to stability, or its causality (Ca); the relative importance of the event in relation to similar previous events or its escalation potential (Es); and the breadth of the stakeholders who are affected by the event, or its centrality (Ce). Thus, the sum (Ca+Es+Ce) depicts the consequence for state instability/stability of an event. This enables researchers to determine event trend lines within the context of individual states while also comparing event trends across states. The first dimension is binary – either the event is positive (stabilizing), or negative (destabilizing). The other three dimensions are evaluated on a three point scale (1 to 3). In sum, Ca+Es+Ce results in a +ve or –ve values range from 3 to 9.

The event is also assessed in relation to two other categories. Events are considered conceptually related to and potentially modifying influences on "structural factors". It follows that events should be classed in one of the six CIFP clusters of security and crime, governance, environment, economic development, demography and human development. In some cases, it was found that clusters were not exclusive and the same event might fall under different clusters although the impact of the event on a cluster might vary.

To summarise, event magnitude was calculated by summing the three dimensions of an event (Ca+Es+Ce) to provide an event score between 3 and 9, with 3 being a relatively minor event, and 9 being highly significant. The event magnitude is then augmented by a +ve or -ve sign to indicate whether the effect has a net decrease or increase in fragility. It is noted that COI determination was based on this score. The analyst used this indicator to explore summary statistics as well as trend lines of the region's events described below. Similar appraisals can be done on events disaggregated among the six indicator clusters; these clusters have been identified as having significant and differential relationships to state fragility. Sorting events by cluster allows the analyst to disaggregate the overall state fragility trend and to differentiate the trend within the various clusters. Finally, all events were classified as challenges

pertaining mostly to either state authority, legitimacy, or capacity, and to a performance cluster whether governance, security and crime, human development, demography, economic, or environment.

When combined with structural data, the summary events data affords a snapshot of a country's baseline and trend performance. Using statistical methods, the CIFP methodology allows a graphical representation of the overall country risk. This general picture can be broken down into six cluster areas, providing a more nuanced rendering. The average event magnitude reflects the mean impact of all events. Positive results are indicative of an environment that, on balance, is experiencing a predilection for stability, as there are either more stabilizing events or more strongly valued stabilizing events. Negative results signify the opposite, an environment characterized by threatening events and a drift towards instability. Magnitude equates to consequence and exposure. The larger the figure, the better (or worse) events effect on state fragility over the given evaluation period. The event trend score is used to observe whether the events demonstrate any positive or negative trend over time. These composite indicators can be plotted against time – usually six months – and trend lines generated, based on ordinary least squares regression.

Typically the trend line is based on a weekly aggregate in order to provide appropriate differentiation and to accommodate the increase or decrease in the total number of events and the associated changes in event magnitude. To attain this trend line, the composite indicators are first summed by week; for instance, if one week has four events with the conflict indicators of +2, +2, -2 and -2, the stabilizing weekly aggregate would be +4, the destabilizing weekly aggregate would be -4, and the overall weekly aggregate would be 0. This weekly average is then plotted over time to produce a trend line to incorporate the belief that an increase or decrease in total number of events should matter in addition to their individual values. That is, one would presume that a rapid increase in the number of stabilizing events would indicate an improving trend, even if the conflict indicators for the individual events remain largely unchanged. This trend analysis provides an overview of the general event developments over the months under consideration. This approach combines systemic scoring with judgement.

Case Selection

The first step in selecting cases entailed reviewing the records representing a diverse range of countries from a number of leading databases. Specifically country indices from four different sources (listed below) were compared:

- 1) FI (Fragility Index) – Country Indicators for Foreign Policy
- 2) SWI (State Weakness Index) – Brookings
- 3) FSI (Failed State Index) – Fund for Peace
- 4) MG (Marshall-Goldstone) - Political Instability Task Force

Although the country indices are highly correlated, the ranking of countries varies which suggests marked differences in the emphasis placed on various indicators of state condition. The values are compared below in **Error! Reference source not found.**

Table 14: Correlation Matrix: 2006 Data (All Countries)

	FI	SWI	FSI	MG
FI	1.00			
SWI	-0.92	1.00		
FSI	-0.88	-0.85	1.00	
MG	-0.90	-0.90	0.81	1.00

Case Selection Criteria

The selection criteria were then applied, as described below, in steps:

Step 1: Select the country

Countries were chosen on the basis of location (global sampling), high FI values, high volatility in FI values, and/or high dispersion among A-L-C values.

Step 2: Identify stable and unstable years

Unstable years were identified based on the highest absolute increase in FI from one year to the next; similarly, stable years were identified based on the highest absolute decline in FI from one year to the next.

Step 3: Confirm data availability for countries and years selected

Data availability was included as selection criteria to allow for comparison of CIFP and VRA datasets. The threshold levels and number of countries for which VRA data was available (1990 to 2006 inclusive) for comparative purposes is indicated below:

More than 2 years VRA data	20 countries
2 or more years VRA data	25 countries
1 or more years VRA data	28 countries

Data Entry

The data were entered into a dashboard in Microsoft Access created specifically for this project. A period of time was assigned to each coder (two years, typically) for stable and unstable periods in a country's history. For every month, a target of 24 events was sought. Each event was entered into the dashboard according to the date the event was reported, the headline from newspaper or magazine reporting the event or a paraphrase of the headline (copyright issue), the publication the reported event came from, the score (+ or -) for stabilizing or destabilizing respectively, and a score of 1 to 3 for each of causality, escalation and centrality. The data was stored in Microsoft Access files, and was converted to Microsoft Excel files as needed to represent significant swaths of data on a single display. A sample of the dashboard is shown in Figure 22.

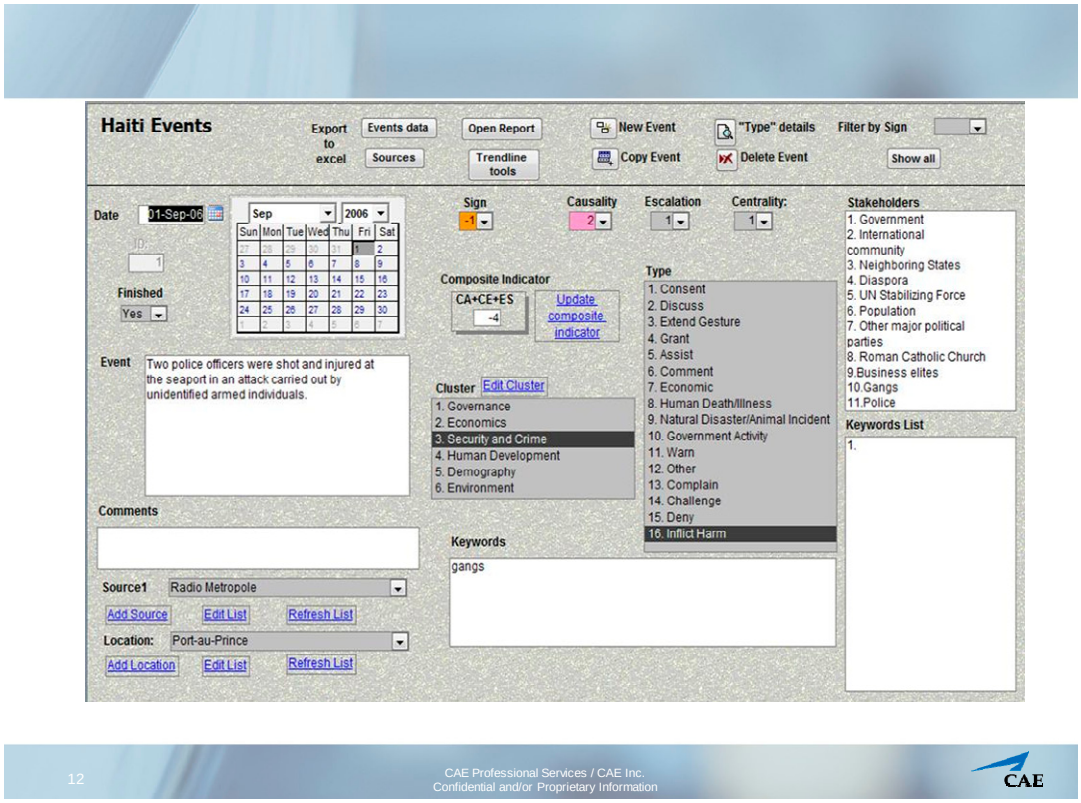

Figure 22: Data Entry Dashboard

VRA Events Coding

The IDEA framework that VRA employs to machine-code its assessments of social, economic, environmental, and political events data "contains summary indicators, such as the coerciveness and contentiousness of events and conflict carrying capacity that can be used to gauge conflict escalation".[46] It is worth noting in passing that VRA applies interstate-biased definitions to intrastate conflict.

The metrics chosen for this work were 1-CS, Hostile, and Goldstein. CS (Country Stability) is a modified conflict carrying capacity measure that assimilates proportions of government and civil direct actions, and all forceful and direct actions. These are characterized using the WEIS categorization. Events are scored if the action involved threat, demonstration, sanction, expulsion, seizure, or force. Hostile action is the proportion to all actions of the WEIS actions stated above excluding threat. The Goldstein scores range from extreme conflict to extreme cooperation, and incorporate all the WEIS-categorized events (Yield, Comment, Consult, Endorse, Promise, Grant, Reward, Agree, Request, Propose, Reject, Accuse, Complain, Deny, Demand, Warn, Threaten, Demonstrate, Sanction, Expel, Seize, Force).

The differences in approach suggest that CIFP events coding captures changes in country stability that might otherwise be missed by VRA events data collection. In order to test this hypothesis, visualization software was used to display and compare the VRA and CIFP results for Yemen 1998-1999, Haiti 1993-1994 and Moldova 1991-1992[47] Each country included at least one period of marked difference between the two datasets, and these periods were qualitatively examined to determine which dataset provided a more accurate picture of the events during that period.

Figure ***23*** is illustrative of an example of CIFP events data graphing. The X-axis plots the timeline by month, while the Y-axis records event scoring. Thus, an upward movement of the line corresponds to a worsening of the country's stability. As indicated in the key in
Figure ***23***, line thickness denotes the number of events that were documented; the thicker line indicates more events than the thinner sections. The line colour provides a further visual reference using colour graduation based on the total sum of all events scored within the month. The dark red portions depict a more de-stabilizing period than lighter red and green sections. The graph also includes reference lines that illustrate the mean and the median, as well as the trend line, which shows the rise or fall in tension trend of the average event score.

[46] Bond, Bond, Oh, Jenkins, Taylor. 2003

[47] "Tableau" is a software application that is based on a technology from Stanford University. It lets you graphically analyze virtually any structured data set, producing charts, graphs, dashboards and reports from a given data set. The software allows the user to customize views, layouts, shapes, colors to highlight patterns, anomalies or anything of note in a given data set.

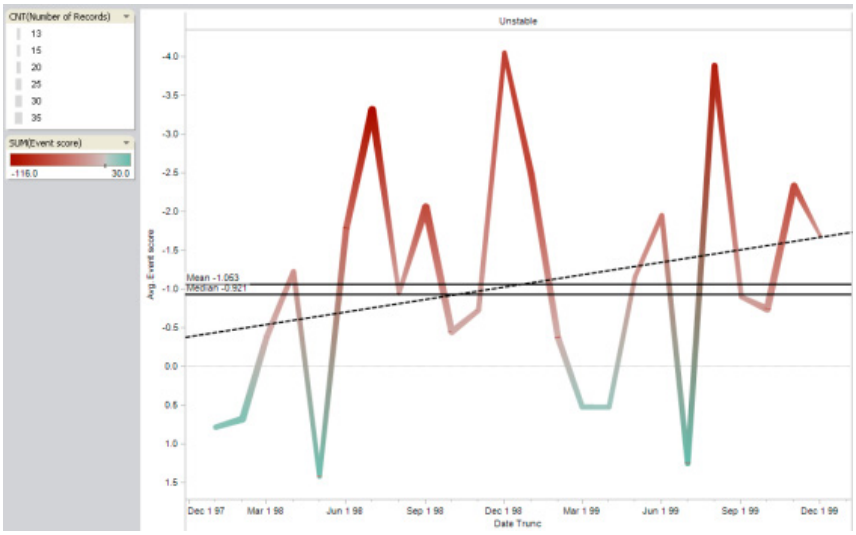

Figure 23: Example of visualization graph

To facilitate a comparison of CIFP and VRA results (using the 1-CS variable), the same time period was used for each country. Hence the differences between the two CIFP and VRA charts relate to the Y-axis scale (event metric in use) and the number of events per month.

Figure ***24*** contrasts the VRA and CIFP events data from Yemen 1998 – 1999. The charts depict similar results for several months, including July 1998, a month which spikes both VRA and CIFP recorded. During this period (circled by ovals), there were numerous explosions and armed clashes with Saudi Arabia over contested island territory. However, in December 1998 (highlighted by rectangles) there is a noticeable difference (see

Figure ***24***). CIFP detected another sharp increase in tension but the VRA results suggest a decline in instability. When examined more closely, this period included multiple explosions as well as a number of hostages killed during a rescue mission. Furthermore, a number of political parties called for a holy war against the United States of America and the United Kingdom. The different outputs using the CIFP and VRA methods show that the data can represent granularity in each event, or aggregate trends over the long term; both are useful depending on the purpose of the analysis.

Figure 24: Yemen 1998 - 1999 Comparison of VRA (1-CS) and CIFP events results

An analysis of Haiti (see
Figure ***25***) produces similar results. The October 1993 protests, violence and deteriorating international relations are captured by both CIFP and VRA, as shown in the oval-circled periods. Conversely, the period highlighted by the rectangles, June 1994, include peaks for CIFP, but not VRA. This month included an increased push for Aristide's return through military intervention. This was construed as destabilizing by CIFP as it created a volatile political environment but was not interpreted and recorded as an unstable period by the VRA methodology, likely because VRA requires more direct action, such as a coup, before political events are considered destabilizing.

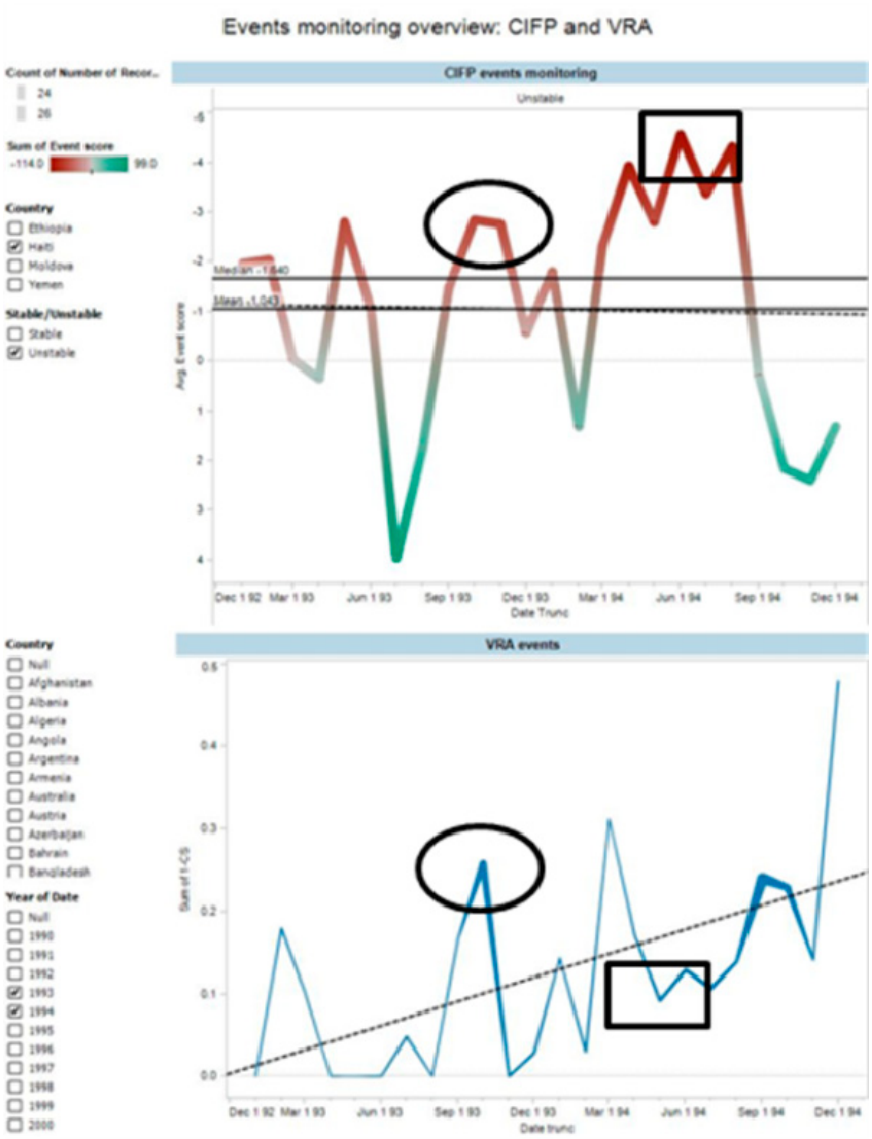

Figure 25: Haiti 1993 - 1994 Comparison of VRA and CIFP events results

The last of the three cases compared was Moldova 1991-1992 (

Figure 26). Again, there is a period during which VRA and CIFP generate similar results. The oval-circled period shows the spike in instability that occurred in July 1992 with the occurrence of major secessionist violence in the Dniestr region. In August 1992, however, there is a significant difference between the two. VRA data would suggest that nothing untoward i.e. impacting on country stability occurred. Conversely the CIFP data notes a drastic increase in stability. An investigation into the contextual setting is instructive and reflects that, during this period, there was an end to the Dniestr conflict and initiation of numerous peace building efforts.

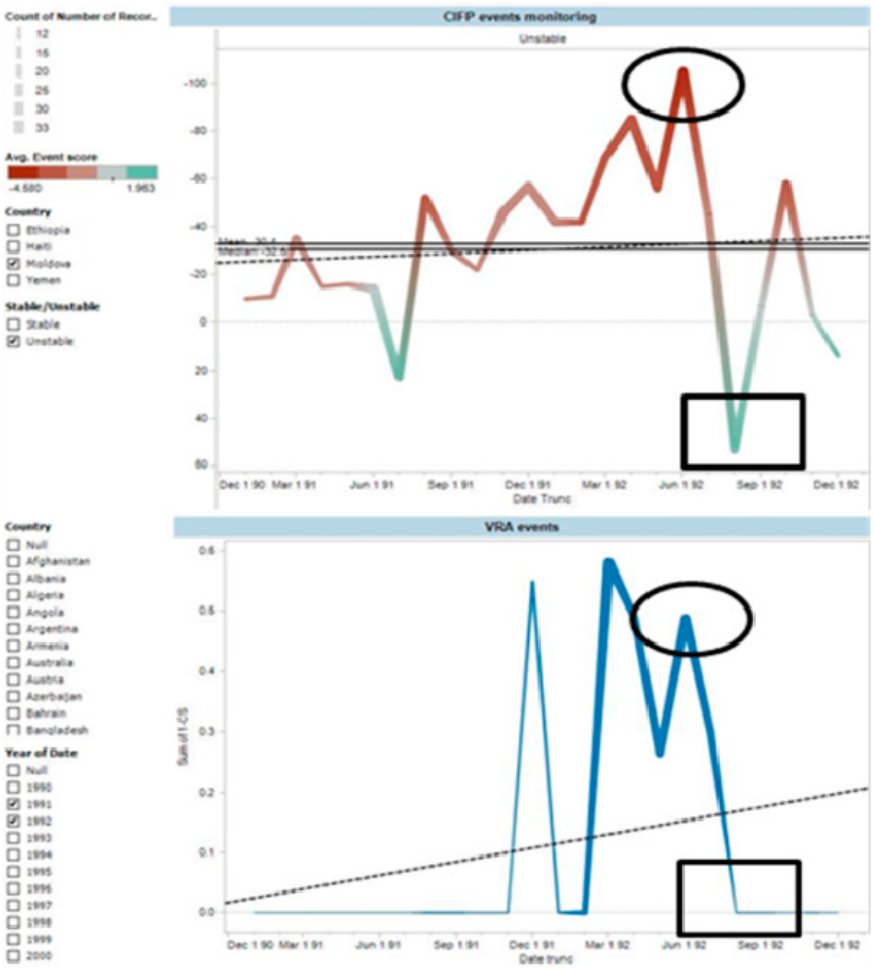

Figure 26: Moldova 1991 - 1992 Comparison of VRA and CIFP events results

This analysis exposes the significant advantages that human coding offers, in particular in relation to the inclusion and interpretation of contextual nuances.

Inter-Rater Reliability

Testing

There are pros and cons to using machine coding and human manual coding when collecting data to analyze. While machine coding follows a rigid algorithm and is reproducible, it cannot account for context, which can provide weight to events. On the other hand, human coding inherently weighs events according to context, but it is subject to coding variance. experiential differences and/or training shortfalls may lead to individual coders assigning dissimilar codes to a similar or even the same event. Taken to an extreme, a database with events coded by multiple analysts could be unusable, due to high dissimilarity between the analysts' coding. This has obvious import if the data being collected is to be used to drive a predictive model to inform policy decisions.

When relying upon human coding, it is important to recognize the validity of the coding results, in order to ensure that the analysts' coding results are sufficiently similar to one another. To determine internal coding consistency inter-rater reliability was assessed as part of this project, i.e. time was taken and a testing process established to ascertain the coders' ability to match results. The coders themselves included both Carleton research assistants, and CAE consultants. All were either Graduates of a Master's level program in International Relations/Political Science or working on degrees in that domain. Initial country assignments exploited unique language skills and/or prior background studies where applicable.

To determine inter-rater reliability, countries (Sri Lanka, Ivory Coast, Chad and Malaysia) were selected from the database. Events from each country were randomly selected and distributed among six analysts, excluding the initial/primary analyst. Each provided secondary coding of all twenty events (five events from four countries). The secondary coders were provided with key coding information, including the country in which the event occurred, the date of the event, a brief description of the event, and a short summary of the context in which the event occurred. A sample of an event put forth for reliability testing is shown below in Figure 27

Country: Malaysia
Date: 12 January 2000
Event: Police arrest opposition members and charge them with sedition
Context: The background to this event centres on the struggle between Prime Minister Mahathir, in power since 1981, and his former deputy Anwar Ibrahim. Ibrahrim was to succeed Mahathir, but Mahathir believed Ibrahim to be too impatient to do so, and he thus expelled Ibrahim who subsequently formed his own party. This period exhibited acrimonious tension between the ruling party (especially the factions loyal to Mahathir) and Anwar's supporters, of which several were former members of the ruling party. Moreover, by the date of this event, Anwar has been incarcerated. The year 2000 began with several crackdowns by Mahathir on Anwar's supporters, and this event involved the arrest of a leading opposition newspaper editor, Anwar's lawyer, and another member of Anwar's party.

Figure 27: Inter-rate Reliability Sample

Using this information, the secondary coders were asked to code each of the twenty sample events and submit their results. These entries were then examined to determine the degree of congruity between the original event coding and subsequent assessments of the same event by secondary coders. Two analysts possessed several years of coding experience across multiple projects. Experience was not determined to be a significant factor; the secondary coders possessed diverse backgrounds in terms of training and practice The other four had less experience coding. Of this latter party, two had been trained individually and the other two had been trained in a group. Of the two independently trained analysts, one had only recently begun coding events, while the other had been involved in a previous project that involved events coding.

Results

Table *13* (below) depicts the results of the inter-rater reliability testing for Malaysia only. It shows the results of 5 different events for all six coders. This table illustrates that, for most events, the discrepancy in how the event was coded is minimal; however, for the more significant events some clear differences arose. For the 13 additional events consult Table 21 in the annex section.

Table 13: Inter-Rate Reliability - Events Compared

Event	CA+CE+ES	Sign	Causality	Centrality	Escalation	Cluster	ALC
14	-8	-1	2	3	3	3	1
14	-6	-1	1	2	3	3	1
14	-7	-1	2	2	3	3	1
14	-6	-1	2	2	2	3	2
14	-3	-1	1	1	1	3	2
14	-4	-1	1	1	2	1	3
15	-3	-1	1	1	1	1	1
15	-3	-1	1	1	1	3	1
15	-3	-1	1	1	1	2	2
15	-4	-1	2	1	1	2	2
15	-3	-1	1	1	1	3	2
15	-5	-1	2	2	1	3	2
16	-5	-1	2	1	2	2	1
16	-7	-1	2	2	3	1	1
16	-6	-1	3	1	2	3	1
16	-6	-1	3	1	2	1	3
16	-6	-1	2	2	2	1	3
16	-5	-1	2	2	1	1	3
17	5	1	2	1	2	2	2
17	5	1	1	2	2	2	2
17	4	1	2	1	1	2	2
17	6	1	2	2	2	2	2
17	5	1	2	1	2	2	2
17	6	1	2	1	3	2	2
18	7	1	2	3	2	1	1
18	6	1	2	3	1	1	1
18	7	1	3	3	1	2	3
18	5	1	1	2	2	1	3
18	7	1	3	2	2	1	3
18	6	1	2	3	1	1	3

Conclusions

These preliminary inter rater reliability tests suggested that there was a higher degree of congruence between those SMEs who were trained collectively than individually. The need for contextual awareness was re-iterated in follow-up discussions with the coders, as it was determined following analysis that

most inconsistencies were attributable to a lack of background knowledge of the country. Nonetheless the divergent scores prompted reflection and discussion.

It is important to note that, inception to date, CIFP training has focused on coding events for a specific country where context, nuance and interpretation are deemed essential in order to track changes in a particular country's performance over time. In design and in intent the data were not intended to be aggregated for cross country comparison purposes and training did not focus on this aspect except in regards to the Ca, Ce and Es entries , their sum and the sign.

The application of an inter-rater reliability test does not assume that entries were necessarily suitable for aggregation. The test simply identifies the closeness of rating among raters, which is important whether the data are aggregated or not. That is, if the intra-class correlation coefficient (ICC) of a particular variable is low between raters, then how is one to judge which rater is correct? How is one to compare assessments of different countries using the same variable if assessed by different raters with low ICC? For example, if Cluster was the variable, then how certain can we be that 'Governance' is the issue for country 'x' by one rater and that 'Security & Crime' is the issue for country 'y' by another rater? Using another example, if a coder determined that a particular event included elements of both authority and legitimacy practice is/was to code as such. Similarly if an event was deemed positively stabilising in the economic cluster but negatively destabilising in the governance cluster then a double entry was made. As noted this additional information, though occasioned in less than 15% of the total number of entries, was deemed essential information for understanding the causal impact of a particular event on a country's stability. As indicated coders were not specifically trained to discriminate within specific events on ALC and Clusters score entries – since these two variables were deemed useful only as background information in the original CIFP events methodology. The original CIFP construct was not designed to utilise either ALC or cluster scoring as aggregated variables for model development nor to use these variables to inform a COI. Hence inter-rater reliability tests were expected to show fair to moderate-level correlations for these variables. On the other hand, a high degree of concurrence among the coders was anticipated when it came to tracking overall instability/stability over time, which was the original intent of the events monitoring system. This is important distinction since the goal was to aggregate the component parts of the overall instability sum, consisting of Causality, Centrality, and Escalation for use in the predictive model. The inter-rater reliability test was accomplished by choosing 18 widely ranging events from conflict to cooperation from various cases amongst the countries examined. The coders were given context on these 'test' cases to assess the events, similar to the procedure when they assessed their separately-assigned cases. The intraclass correlation coefficient (ICC) is the appropriate test statistic for these data.[48] Values less than 0.2 are considered slight, between 0.21 and 0.4 are fair, between 0.41 and 0.6 are moderate, between 0.61 and 0.8 are substantial, and above 0.81 are near perfect.[49] ICC values for the 18 test events were 0.105 for ALC discrimination, 0.389 for cluster discrimination, 0.360 for Causality scoring, 0.472 for Centrality scoring, 0.288 for Escalation scoring and 0.806 for overall instability scoring.

The results are consistent with expectations and, despite the fair to moderate reliability among the Causality, Centrality, and Escalation scoring; a remarkably high reliability was attained for the sum of these component parts. The discrepancy in coding accuracy between the component parts and their sum may be due in large part to a scale effect in which the 3 point scale are susceptible to undue emphasis on

[48] Walter, S.D.; Eliasziw, Michale; Donner, Allan. *Sample Size and Optimal Designs for Reliability Studies.* Statistics in Medicine. Vol 17, Issue 1. Pp101-110

[49] Donner, Allan; Eliasziw, Michael. *Sample Size Requirements for Reliability Studies.* Statistics in Medicine. Vol 6, Issue 4. Pp441-448.

differences rather than similarities and which is not the case for the summary scale which is measure from 3-9 points. Several articles support of this conclusion.[50]

In brief inter-rater reliability proved a useful exercise. Should aggregation of all coded entries for all variables become a project goal, then future development on the coding side should take into account the need to modify the individual components of the summary scale to reduce scaling effects. There should also be efforts to streamline and clearly operationalize entries for both ALC and cluster variables so that overlap between categories is eliminated. One very obvious way to ensure higher ICC scores on variables is to ensure coders have a clearly defined and exhaustive set of categories and to run pre and post test training until such time there is a clearly understood demarcation between the categories. In summary, the results indicate that human coding can be reliable with appropriate training and contextual understanding.

4.4 Dependent Variable and Crisis of Interest (COI)

Overview

The analytical challenge lay in identifying a Crisis of Interest (COI) and determining the dependent variable. It was agreed that a COI reflects a fundamental challenge to state stability, i.e. a significant contesting of state authority, capacity and/or legitimacy. Conceptually this occurs most commonly and is momentous when there is a pre-existing level of tension in the state and it takes the form of political instability and armed conflict, but a COI may also appear in the form of tests of governance and economic performance. The dependent variable chosen was the country-half month. When a country-half month contained a crisis of interest, the dependent variable is 1, if not 0. The choice of a half-monthly period reflects a conscious attempt to maximize the number of data points while ensuring a statistically meaningful number of events were included within each period. Calendar months were used and the data arranged in half-monthly periods beginning with the first 15 days of a month and the balance of the month for the second half.

An initial analysis of the data indicated that the majority of COI involved challenges within and to authority structures and, more rarely, legitimacy and capacity structures.

Criteria for Identifying an Event as a COI

A first step to identifying symptomatic patterns and precursors to state failure was to establish what constitutes a Conflict of Interest. Even with a wealth of data and the advantage of *post hoc* study recognizing and defining a COI proved challenging and highlighted the importance of "equifinality" - there may be multiple paths to the same end.[51] The criteria initially considered included:

1. A change in actors – an increase in the number of actors and/or change in the kinds of actors engaged in or affected by the crisis (e.g., targeting of civilians which previously were not part of the crisis and possibly leading to reciprocal and escalatory behaviour, emergence of paramilitary

[50] Lilly Neumann "Effects of scale length on means and correlation coefficients" *Quality and Quantity* (Vol. 17, no. 5, 1983). Warren Martin, "The effects of scaling on the Correlation Coefficient A Test of Validity (193) Journal of marketing Research Vol. X.

[51] Rohloff, Kurt; Asal, Victor. *Computational Models to Discover Sets of Patterns of Behaviours that Precede Political Events.* Association for the Advancement of Artificial Intelligence: 2008.

forces, a coup leading to military dictatorship, intervention by a military force, the destruction of an ethnic group).

2. A change in issues – a change or broadening of the issues of interest which were the original basis for interactions between actors. These can be understood as economic, political, social, military-security (e.g., an economic grievance leads to political challenges from a particular group or a social issue - is redefined as a security threat).
3. A change in rules – a change in the procedures, rules and norms (explicit or implicit) that guide and inform interactions between actors. These may be either informal or formal procedures, rules and/or norms (e.g., severe human rights violations, or severe curtailment of movement such as martial law or curfew), a suspension in the rights of opposition leaders, collapse of peace talks, collapse of a parliamentary system, declarations of secession etc).
4. A change in structure – a change in the relative distribution of capabilities and/or capacities among actors (defined broadly as power but understood as having economic, political, military-security and social components). Capability can be viewed as a state of development (knowledge, skills) whereas 'Capacity' can be viewed as a 'reserve' for countering shocks and deploying capability. Both define structure.

Hence it was determined that a COI is identifiable as a specific event that embodies the qualitative changes described above within a half-month period. Quantitatively, however, the criteria must be tailored specific to the country analyzed since the level of violence can vary considerably amongst countries to create the same destabilizing effect. Hence, a COI is selected IF within a half-month period:

1. An event scores -9 on the events monitoring scale (where the centrality, escalation and causality dimensions are each coded -3, the highest possible negative score) ;
2. The percentage of destabilizing events (i.e., < 0) exceeds a certain threshold (say 'x') and/or the percentage of highly destabilizing events (i.e., < -5) exceeds another threshold (say 'y').

More restrictive criteria may include:

3. More than one -9 event within a half-month.
4. The event is coded under the 'Authority' dimension and/or 'Security and Crime' cluster.

The original estimate was that using the less restrictive criteria would involve about 60% of the country cases. Using the more restrictive criteria would reduce the sample down to 30-40%.

5 Analysis

This section presents the results of examining events and structural data, and comparing structural data between the CIFP Fragility Index and PITF. Planned comparisons for eventual model testing are based on a 2x2 crossover design to establish the best fit to the statistical model. VRA machine-coded (section 4.3.2) and CIFP's human-coded (section 4.3.3) events data will be combined with PITF's (section 4.2.1) and CIFP's (section 4.2.3) structural factors for this purpose.

5.1 Events and Structural Data – Preliminary Findings

Key objectives of the project were to examine events data, the relationship between structural and events data, and the feasibility of integrating structural and events data to measure state tension. Preliminary analysis found that Crises of Interest were present in stable periods or not visible enough in unstable periods, making it difficult to distinguish periods of stability and instability in the modelling and regression analysis. For example, the relative scarcity of -9 events, especially during unstable periods, suggests that instability does not necessarily manifest itself in an extreme fashion. As pointed out earlier, the use of the terms "unstable" (=U) and "stable" (=S) refers to years during which the biggest deteriorations or improvements (in an absolute sense) in fragility scores were recorded. E-U refers to events(=E) in an unstable (=U) year; (-6,-9)-U refers to all events with a score ranging from -6 to -9 in an unstable year; F-U refers to fragility scores in an unstable year.

Two (expected) results stand out from this analysis:

- First, the average and volatility of COI for unstable years is higher than for stable years (16.65 > 7.15 and 13.23 > 6,25 respectively; a t-test of differences in means rejected the null hypothesis that the two means are equal[52]). With exceptions for Albania and Chad, the percentage of COI in unstable and stable years was higher, as anticipated. Thus, the years that were chosen reasonably reflect "periods of relative instability and stability" as discussed in the section on case selection. In the case of Albania and Chad, lower-intensity events compensated so much that they gave rise to the unexpected negative results.
- Second, a higher percentage of COI are correlated with higher fragility scores for stable and non-significantly for unstable years (see **Error! Reference source not found.** and 29 below).

[52] This is especially important if the intention is to examine statistical significance across groups.

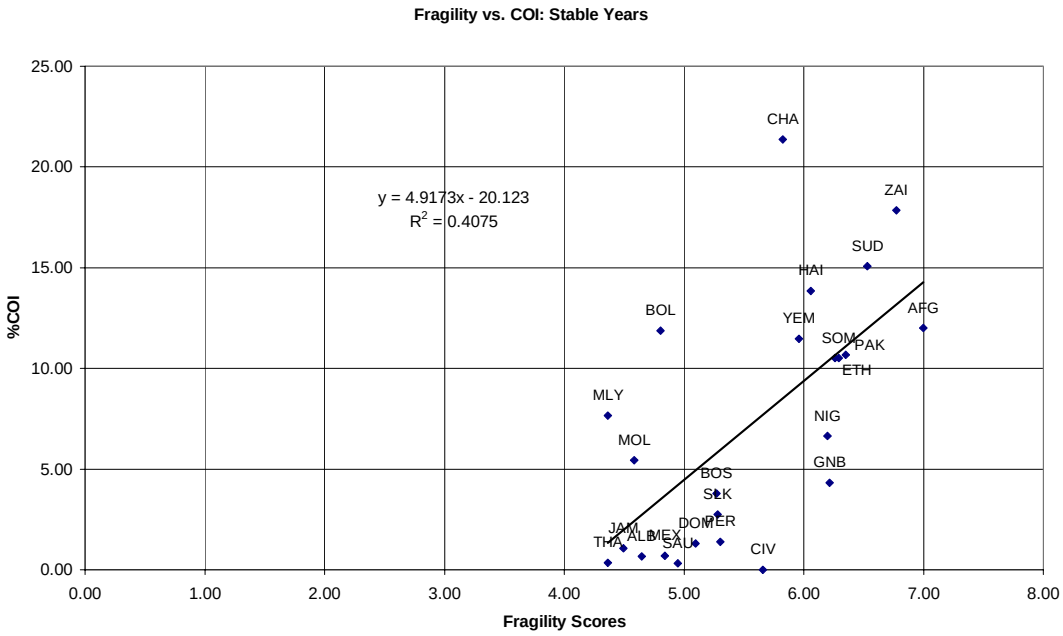

Figure 28: Structural and Events Data, Stable Years

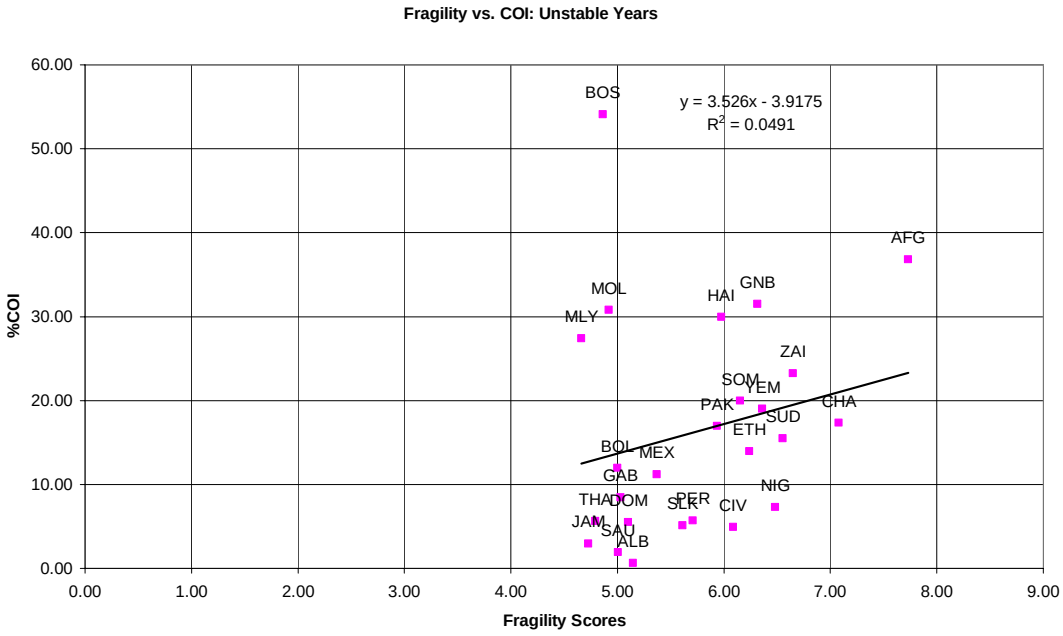

Figure 29: Structural and Events Data, Unstable Years

The ALC framework was also exploited to investigate how authority, legitimacy and capacity scores changed during the years that were chosen as stable or unstable periods. In most cases, changes in authority (50% and 79% of cases in stable and unstable periods) were the main drivers of improvements or deteriorations in the fragility scores. Furthermore, the highest percentages of COI were related to changes in authority, for both stable and unstable periods.

A correspondence between structural and events data was expected. These data may provide quite different insights into what is happening in individual countries. In the case of Ethiopia, for which 1994 was identified as an unstable year based on the change in the fragility score as depicted in Figure ***30***. This trend continued even when additional years of data were added (not shown here). The trend line for events data also shows an increase in fragility (recall that –ve event scores denote increased de-stabilization) but there are fluctuations in the monthly data, providing more nuance to the analysis than structural data can only depict. Ultimately, it is this type of nuance that the model calibration can capture.

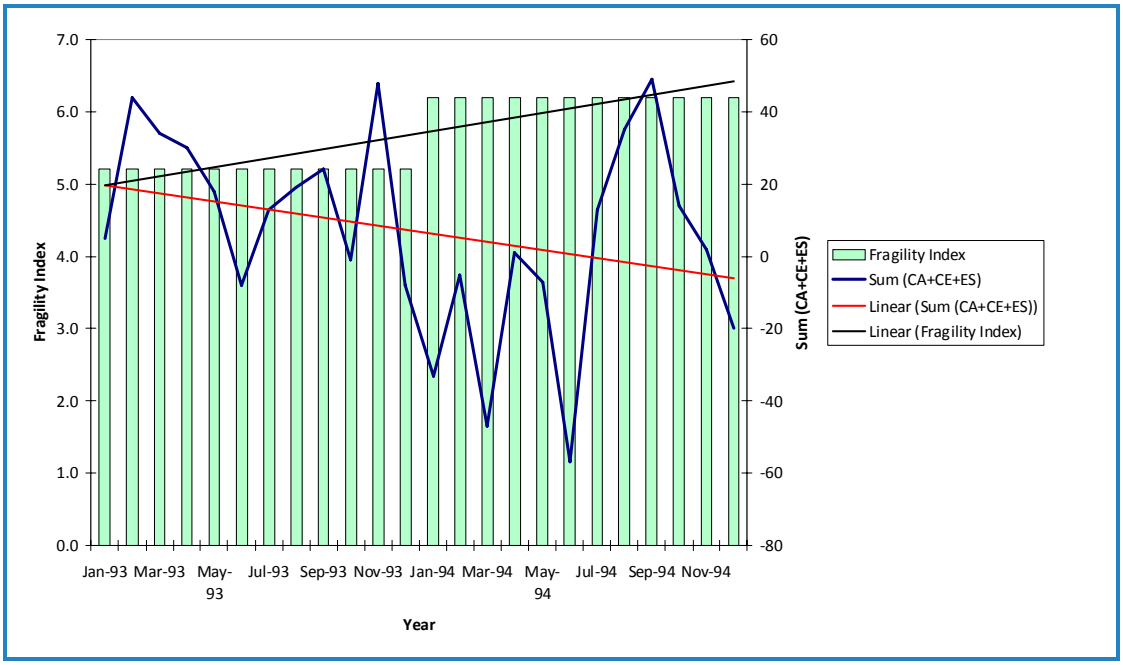

Figure 30: Structural vs. Events Data, Ethiopia

5.2 Comparison Between CIFP's FI and PITF Structural Factors

PITF data were not available for all the years of interest for the following states: Saudi Arabia (2004), Chad (2004-5), Guinea-Bissau (2004), Moldova (1990), Bosnia (1990-1), and Jamaica (2004-6), and were thus excluded from the comparison with FI. A table showing the comparison between stable and unstable groups is shown in Table 14. Neither FI nor PITF were found to be significantly different ($p < 0.05$) between stable and unstable years for the remaining 19 states. However, FI and PITF were found to be significantly correlated ($r = 0.42$). Perhaps one should not be surprised by the lack of significance between stable and unstable periods because of the way that these periods are defined (namely that an extremely weak state, in our definition, may be characterized by periods of relative stability, and vice versa). Caution should also be exercised in comparing FI and PITF data since these two indices measure different things and are used differently; furthermore, only the FI was used to select stable and unstable periods, as defined above. This point is reinforced in the descriptive statistics provided below. We should not be surprised that the statistics being reported for unstable and stable when matched against VRA and PITF data are not significant since the CIFP operationalization of stability and instability is distinct from that used in either of these projects. It is somewhat akin to trying to fit a square peg in a round hole.

Table 14: Stability T-Test Comparison

T-tests: STABLE (Group 1) vs. UNSTABLE (GROUP 2) Years											
	Mean 1	**Mean 2**	**t-value**	**df**	**p**	**N 1**	**N 2**	**SD 1**	**SD 2**	**F-ratio**	**p**
PITF	-3.47	-3.58	0.55	36	0.58	19	19	0.70	0.60	1.37	0.51
FI	5.86	5.41	1.57	36	0.13	19	19	0.95	0.78	1.50	0.39

Descriptive Statistics of VRA Events Data

VRA events data were incomplete (i.e., there was no record of events during certain monthly periods) for 9 states. The analysis below was conducted on the remaining 16 states: Peru, Nigeria, Saudi Arabia, Ivory Coast, Albania, Sri Lanka, Haiti, Yemen, Malaysia, Pakistan, Zaire, Afghanistan, Ethiopia, Sudan, Mexico, and Thailand. None of the conflictive/co-operative metrics were found to be significantly different between the stable and unstable years, shown in **Error! Reference source not found.**

Table 15: VRA T-Test

T-tests: STABLE (Group 1) vs. UNSTABLE (GROUP 2) Years											
	Mean 1	**Mean 2**	**t-value**	**df**	**p**	**N 1**	**N 2**	**SD 1**	**SD 2**	**F-ratio**	**p**
1 - CS	0.18	0.23	-0.92	30	0.36	16	16	0.17	0.14	1.57	0.39
Hostile	0.25	0.24	0.07	30	0.94	16	16	0.19	0.16	1.42	0.51
Goldstein	-0.93	-1.39	0.68	30	0.50	16	16	1.83	1.99	1.17	0.76

Descriptive Statistics of CIFP Events Data

A total of 26,277 events were scored for 25 countries over 4 years. The designations beginning with 'O' refer to the overall results, 'A, L, and C' refer to institutional challenges to Authority, Legitimacy, and Capacity, and 'G, M, S, H, D, E' refer to performance clusters of Governance, Economy, Security and Crime, Human Development, Demographics, and Environment. 'O_NUM' refers to the number of events scored, which was found to be higher by about 1 event/month in the unstable vs. stable years. '% all' refers to the percentage of events attributed to an institutional challenge (i.e. A, L, or C) or to a performance cluster (i.e. G, M, S, H, D, or E). The percentage of events was higher during the unstable vs. stable years for 'L and S', and lower for 'G'.

'AVG' and 'SD' refer to the average and standard deviation of the events' scores. These values were respectively found to be more de-stabilizing (i.e. lower in value) and more volatile (i.e. variance) during the unstable vs. stable years for 'O, L, G, and S'. '-9' and '%-9' refer to the average number and percentage of -9 events during a month, which were significantly higher during the unstable vs. stable years for 'L and G'. '%<0' and '%<-5' respectively refer to the percentages of events that were negative (i.e. de-stabilizing) and less than -5 (i.e. highly de-stabilizing); both were significantly higher during the unstable vs. stable years for 'O, L, G, and S'. '%<-5', but not '%<0', was significantly higher during the unstable vs. stable years for 'A, C, M, and H'. Please refer to Table 22 in the annex

Statistical Comparison Between CIFP and VRA Events Data

Correlations with the CIFP events data tend to be highest for the Goldstein metric, which encompasses the complete range of conflictive and co-operative events. The highest correlations for all three VRA metrics were found for 'A_%<0' (percentage of de-stabilizing events that were challenges to authority. The results are shown in the annex in ***Table 23***.

5.3 Conclusions

The events analyzed by CIFP and reported herein represent a uniquely valuable dataset for calibrating the conceptual model of early warning of a COI. The challenge will be to refine the definition of a COI such that it is not so rare to avoid difficulties with regression, yet not too frequent so that the COI is sufficiently meaningful to warrant attention. For example, '-9' events are deemed most de-stabilizing, but they occurred in only 11 of the 25 states examined. Though it is important to note a greater percentage of COIs occurred during unstable years than stable years. While not surprising this is the kind of brush clearing, largely impressionistic research that will be required for the project to develop meaningful results. As was discussed in section 5.2, we relaxed the criteria for COI events as ones ranging in score from -6 to -9 to circumvent the problem of too few '-9' events. Hence, COI criteria that consider less, but still highly de-stabilizing events are required for the majority of states examined herein. Of those that experienced '-9' events, Yemen, Moldova, Zaire, Sudan, and Somalia had eight or more during the four years analyzed. A more restrictive definition of a COI beyond just '-9' events might be necessary for these cases and our research in the long-term should examine how our results are sensitive to the restrictions that we place when defining a COI (as explained earlier).

The analysis of events data leading up to a COI, as identified by PITF, was the initial task specification. However, strong arguments were made against taking such a narrow approach. It is important to keep in mind that the PITF adopts a conflict-based definition of state failure and focuses on severe cases of state failure but not the full range of fragile states that CIFP examines. The latter has the benefit of allowing the identification of earlier turning points, entry, as well as sequencing and timing. Hence, 2-year periods of stability and instability were identified for events analysis and it was subsequently found that COIs (defined by the occurrence of highly de-stabilizing events) occurred at various times and were not confined to just the unstable period. Unfortunately, this placed a limitation on the period of time that events could be analyzed prior to a COI. To alleviate this constraint, it is recommended that additional events data of at least 6 months are analyzed prior to the year of interest.

Preliminary exploration of the data using logistic regression found that structural factors by themselves did not contribute significantly to the prediction of a COI because COIs occurred in both stable and unstable years, which precluded any possible discrimination with the choice of structural factors applied herein. –Notwithstanding this result, it is important to know in advance which structural factors are relevant and which are not. As we did not run a simple model or conduct a factor analysis specifying the structural causes of failure, we can only guess what those factors might be. Fortunately there are a number of structural models of state failure and fragility we can draw on including that developed by Carment et al (2009) that would be a source for us to draw on.

Events are deemed relevant \ because COIs are dispersed in the various years examined, thus reducing any uniqueness in the state's baseline condition either in the year or the year before a COI occurred. Therefore, if structural factors are to contribute significantly to the LR, they must be known more precisely than just annually and closer to the occurrence of a COI. Or, periods in a state's history must be

chosen that are devoid of COIs (i.e., truly stable), which should also be reflected by a more stable baseline condition that is markedly different from that state's unstable period involving COIs. Precision is important, and closeness raises a number of difficulties because, by definition, structural data will always be indexed annually.

There are a number of relevant conclusions that can be made after reviewing the ambitious agenda and the results produced thus far.

In attempting to clarify similarities and differences among four different datatsets we can be firm in concluding that they are each asking distinct questions, have different dependent variables and different ways of measuring failure and instability. As a result we should not be surprised by the lack of uniformity between these data sets and the lack of correspondence in identifying periods of instability. In fact VRA and PITF were never intended to be combined to together as one is dealing with conflict and the other failure. Conversely the CIFP events and structural dataset draws from the same theoretical basis making synthesis a more reasonable prospect. In brief, there are differences in the databases such that they are trying to measure different things, leaving comparative analysis limited. Exacerbating this challenge further, there is also no universally accepted definition of stability/instability or what constitutes a failed or fragile state.

We also found that more COIs happen during years of instability, a promising result that was expected. This would lead us to conclude that we should focus more on deepening our understanding of why COIs occur during periods of instability rather expecting them to occur during periods specified by PITF as "failure" period's. Indeed it should be a prime motivation of this research to strive to understand the kinds of structural features unstable states face when compared to their more stable counterparts in order to pinpoint the precursors to COIs. We can with sufficient thought find corresponding matches between events and structure as they are organized around the six cluster areas and ALC attributes.

We also conclude that with proper training, this project can achieve higher inter-coder reliability. However, one needs to be aware beforehand of what the events data will be used for, and the level of differentiation that is required for each specific measure, and training provided to coders accordingly. The advantage of the CIFP/SME event coding procedure is evidenced in comparison to VRA. The greater detail, as compared to the latter warrants the additional cost to achieve clarity and precision and eventually, real time country monitoring.

The wealth of data analyzed merits significant additional analysis. Much of this should and can be in form of broad brush stroke work to explore patterns (e.g., inductive pattern recognition, process tracing, incremental stability changes, correlational analysis). In this regard the CIFP Tables of countries ordered by A, L and C, and by the six cluster areas can provide instructive insights in to how to work with and develop a typology of state failure.

6 Recommendations

The immediate work ahead appears to be focussed heavily on calibration and validation of the EWM using the events data collected in this study. Macro-structural models though insufficient in themselves (since their forecast horizon is typically 2-5 years) could be developed as a parallel process in order for us to develop a typology and to determine the conditions under which an unstable state experiences a COI. This information could feed into the formal model for calibration purposes. Though the formal model relies upon events data for the identification and subsequent prediction of COIs a refined structural data set must be better utilised given that we now know that the majority of COIs constitute challenges to authority and fall within the security and governance clusters during unstable periods. As O'Brien observes" integrated crisis early warning is fundamentally concerned with identifying those perhaps seemingly benign, policy relevant factors that, when combined with other factors, systemically preceded crises in a probabilistic way".[53] If one were to develop a long, rather than short term perspective, identifying the array of structural factors associated with COIs may help policy makers in developing long term prevention policies to immunize states against potential future challenges and structural shortfalls. Further, it is envisioned that the selection criteria of COIs must be country specific and context sensitive and having detailed structural profiles of these countries will help in that selection.

To address these challenges, a list of known COIs that serve as models for testing could be developed. Then, a set of key indexed structural factors that have led to their development, stakeholder analysis, a list of trends reporting where key conditions are currently present and inventory of preventive tools could be assembled. By processing this information in a systematic fashion, it might even be possible to forecast which crises are likely to generate failure. The procedure could be as follows:

1. Identify and create several models of known types of crises (COIs) that are associated with instability – e.g. humanitarian, political violence, ethnic conflict etc;

2. Determine the probability of crisis occurrence using CIFP's fragility structural data based on these models comparing stable and unstable periods. A co-relational framework would be developed;

3. Augment the structural data analysis for each type of crisis occurring in each country with additional information drawn from initially the events data, Trigger and decision maker analysis data could be added at a later stage to refine the model even further to generate greater precision;

4. Regression procedures could provide a practical form of linking the incidence of a given type of crisis (definitionally, a zero one variable) to the myriad of underlying factors including events, etc

A 4-stage process is suggested, beginning with an assumption of Crisis Occurrence = constant + $\beta_k \cdot X_i$ + random $error_i$ where 'i' is the number of observations and 'k' is the number of indices.

Lagged variables will be included where appropriate. The specification of logit regression models will involve a number of trials before arriving at models with sufficient explanatory power. The degree of explanatory power would be judged by comparing true positives, true negatives, false positives, and false negatives of the predictions of crisis situations that a given set of independent variables produce in a least squares framework.

[53] Sean O'Brien. Crisis Early Warning and Decision Support; Contemporary Approaches and Thoughts on Future Research" International Studies Review (2010) 12, p98.

For stage 2, we would generate country profiles after estimating an empirical model based on past data. By substituting current country variables into the given regressions, the estimated probabilities of crisis occurrence—the degree of risk—using an inventory of specific types of crises for every country would be generated. Producing specific models to be applied against specific crises, and the conditions under which triggers are effective in fomenting conflict, in addition to leadership models influencing the action and reaction of organizations.

In stage 3, dynamic data – e.g. events and stakeholder interests would be included in the country profiles, to bring greater precision to specific country analysis.

At stage 4, in the prevention domain, the coefficients of prevention and management tools variables may lead to a recommendation based on the strategies that have worked most effectively in the past at averting similar situations. This inventory will be matched against the country profiles.

To meet this goal, it is recommended that in-depth events data collection be undertaken for more countries over a greater period of time and that this be applied in real time as an events monitoring tool for decision support. While large, the current events data set must be larger and ongoing in order to satisfy requirements of statistical accuracy. An improved data set may also provide more nuance to the current analysis and decision making. If successful, then the diplomatic and intelligence communities might consider augmenting the model's predictive power with classified data. Decision support must complement existing processes and acknowledge decision cycles, and analysis must recognize the capabilities and resources that a decision maker has at their disposal. Providing active real time monitoring in conjunction with structural data can help ensure those resources and capabilities are exploited more effectively and allocated in a timely manner.

A second and related recommendation is to continue to refine the term COI, drawing on the research conducted on triggers in Phase I of the project.[54] It is suggested that a balance between identifying the causes of rare events, on the one hand, and ensuring statistical accuracy on the other be determined. This can be done through a detailed qualitative assessment of known COIs to ascertain exactly what the component parts of a COI are. For example, it has been established that most COIs constitute challenge to state authority and fall within the security and governance clusters during unstable periods. For each country, attention should now focus on changes in these clusters in order to find out impressionistically, qualitatively and statistically what the structural conditions are that generate events leading to COIs, and whether some general patterns may be observed from the relationship between structural conditions and COIs. It is recommended that the Gurr/Moore tripartite risk assessment methodology[55] alluded to previously be applied in future stages of analysis. This would entail regression based models using structural data only, the creation of risk profiles of type of state failure using CIFP to data to compare outcomes and theoretical regressions.

A third recommendation is to differentiate, or at least refine, our understanding of instability and state failure. Greater clarity might be brought to this inquiry by categorizing stable and unstable periods according to the value of the fragility index, whether it is changing, and if so in what direction, as proposed in the following table. Implicit in this taxonomy is that states are stable if FI is low, but that

[54] Hales, Doug; Miller, Jordan; Tikuisis, Peter. *Triggers of State Failure.* CAE Professional Services/DRDC Toronto-Adversarial Intent Section: 2008. CR 2008-054.

[55] Moore, W; Gurr, TR. *"Assessing risks of ethnopolitical rebellion in the year 2000: three empirical approaches'* in Schmeidl & Adelman. Synergy in Early Warning Conference Proceedings. September 1998.

changes in FI allow a more nuanced interpretation (e.g., states at Stable Level III might have more in common to states at Unstable Level I than states at lower Stable Levels).

Table 16: Proposed categorization schema for defining stable and unstable periods.

Value of FI	Change in FI		
	Decreasing	Steady	Increasing
Low	Stable Level I	Stable Level II	Stable Level III
High	Untable Level I	Untable Level II	Untable Level III

The PITF coding in which a state is either failed or not provides little interpretative power since state failure remains a rare occurrence. Broadening our understanding/interpretation of failure (tapping into perhaps higher thresholds of fragility as measured by CIFP for example instead of failure) or extending our evaluation to other dimensions of failure and the relationship of COI to them (such as low capacity or loss of legitimacy) allows for conditions that may occur in the absence of large scale violence. These objectives can be realized by increasing the period and number of countries for which data are collected, recognizing that there are constraints on how far back one can access events data. The period 1990 onwards does not seem unreasonable.

A fourth and final recommendation is to disseminate the information, methodology and findings gleaned from this project to the broader academic and defence research communities. The outputs from this project should become the basis for submissions to peer reviewed journals and high quality international conferences. The research team should seek out new collaborators where warranted and seek input from expert opinion as much as possible. Opportunities should be sought to sustain and extend the discussion and collaboration.

7 Further Research & Collaboration

There are a number of opportunities for collaborative research and venues for building on the current work in events data analysis. The parallel DARPA initiative under the direction of Sean O'Brien, for example, is currently developing an Integrated Crisis Early Warning System (ICEWS).[56] Its task is to exploit a capability-based approach to crisis response strategies for the US government. The initiative unfolds in three phases: i) forecasting models are developed, ii) real time monitoring capabilities are deployed and, iii) in-theatre capabilities are applied to provide support to Combat Command. The DRDC project can also evolve at least through the second phase to then provide decision support to CF Development and its Future Security arm. Beyond the use of structural and events based data analysis, some elements of the project signal a direction in which this research project could move including incorporation of data on leadership capabilities and styles. It is noteworthy that O/Brien et al[57]. also recommend taking advantage of the latest in information technologies to amass and analyze data.

Michael Ward's work[58], alongside the research of Philip Schrodt[59] [60] [61], figures prominently in the DARPA initiative. His use of text parsing data analysis tools to predict conflict in Africa draws on insights similar to our team's efforts. Structural data by itself is insufficient to draw firm conclusions about the intensity and frequency of intra-state conflict. What is required is more nuanced and detailed information regarding the date and location of conflict-related events. Such information is difficult to assemble on a large scale, though the CIFP methodology certainly provides a means for doing it more succinctly than that used in Ward's Armed Conflict and Location and Events Dataset (ACLED)[62].

Interesting research is also being conducted by the European Union and the World Food Programme, both of which are tapping into events monitoring capabilities for crisis response. This includes work by Rohloff and Assal[63] which uses pattern recognition as a technique for identifying Events of Interest (EOI), a point taken up by Stoll, Trappl[64] and others who have worked extensively on neural networks and pattern recognition for crisis early warning. In terms of recommendations it would be wise to continue to track these initiatives and draw on their research where appropriate. In particular, all of these projects have similar goals and objectives to the extent they are looking to devise a real-time capability sufficiently robust that it can be used as a decision support tool. Clearly this is an important and topical area of research which deserves continued investment and collaboration from the Government of Canada.

---

[56] O'Brien, Sean P. "Crisis Early Warning and Decision Support: Contemporary Approaches and Thoughts on Future Research" in *International Studies Review* (2010) 12, 87-104.
[57] Ibid.
[58] Walter, S.D.; Eliasziw, Michale; Donner, Allan. *Sample Size and Optimal Designs for Reliability Studies.* Statistics in Medicine. Vol 17, Issue 1. Pp101-110
[59] Schrodt, P. A. (1999) "Early Warning of Conflict in Southern Lebanon using Hidden Markov Models," in Starr, H. (ed.) *The Understanding and Management of Global Violence*, New York: St. Martin's Press.
[60] Schrodt, P. A. (2000) "Pattern Recognition of International Crises using Hidden Markov Models," in Richards, D. (ed.) *Political Complexity: Nonlinear Models of Politics*,. Ann Arbor: University of Michigan Press.
[61] Schrodt, P. A., and Deborah J. Gerner. (2000). "Cluster-Based Early Warning Indicators for Political Change in the Contemporary Levant." *American Political Science Review* 94,4.
[62] Available online at: http://www.acleddata.com/
[63] Rohloff, Kurt; Asal, Victor. *Computational Models to Discover Sets of Patterns of Behaviours that Precede Political Events.* Association for the Advancement of Artificial Intelligence: 2008.

[64] Trappl, Robert (ed). *Advances in Group Decision and Negotiation – Programming for Peace: Computer-Aided Methods for International Conflict Resolution and Prevention.* Dordecht, Netherlands: 2006

8 Annexes

Appendix A CIFP – Data Sources (see www.carleton.ca/cifp for details).

1. Governance
Freedom of the Press*
Government Effectiveness†
Level of Corruption ‡
Level of Democracy§
Level of participation in international political organisations**
Percentage of Female Parliamentarians††
Permanence of Regime Type§
Refugees hosted ‡‡
Restrictions on Civil Liberties*
Restrictions on Political Rights*
Rule of Law†
Voice and Accountability in Decisionmaking†
2. Economics
Economic growth -- Percentage of GDP††
Economic Size -- Relative -- GDP per capita††
Economic Size -- Total -- GDP ††
External Debt -- percentage of GNI††
FDI -- percentage of GDP††
Foreign Aid -- percentage of Central Government Expenditures††
Foreign Aid -- Total per capita††
Inequality -- GINI Coefficient††
Inflation††
Informal Economy -- Black Market§§

Table 17: Cluster List - Governance and Economics

Informal Economy -- Ratio of PPP to GDP[††]
Infrastructure -- Reliability of Electricity Supply[††]
Infrastructure -- Telephone mainlines per capita[††]
Internet Usage per capita[††]
Investment Climate -- Contract Regulation[§§]
Level of participation in international economic organisations[**]
Paying Taxes[***]
Regulatory Quality[†]
Remittances Received -- percentage of GDP[††]
Reserve Holdings – Total[††]
Trade Balance -- percentage of GDP[††]
Trade Openness -- percentage of GDP[††]
Unemployment – Total[††]
Women in the labour force[††]
3. Security & Crime
Conflict intensity[†††]
Dependence on External Military Support[‡‡‡]
Human Rights – Empowerment[§§§]
Human Rights -- Physical Integrity[§§§]
Military Expenditure -- percentage of GDP[††]
Political Stability[†]
Refugees Produced[††]
Risk of ethnic Rebellion[****]
Terrorism -- Number of fatalities[††††,‡‡‡‡]
Terrorism -- Number of Incidents[††††,‡‡‡‡]

Table 18: Cluster List - Security & Crime

4. Human Development
Access to Improved Water††
Access to Sanitation††
Education -- Primary Completion -- female††
Education -- Primary Completion -- total††
Education Primary Enrolment -- total††
Education -- Primary Enrolment -- Ratio of Female to Male††
Food Security -- Aid as percentage of total consumption§§§§
Gender Empowerment Measure*****
Gender-related Development Index*****
Health Infrastructure -- Expenditures as a percentage of GDP††
HIV/AIDS -- New AIDS Cases Reported†††††
HIV/AIDS -- Percentage of Adult Females Infected††
HIV/AIDS -- Proportion of Adult population infected††
Human Development Index*****
Infant Mortality††
Literacy – total††
Literacy – female††

Table 19: Cluster List - Human Development

5. Demography
Life Expectancy – Female††
Life Expectancy – Total††
Migration -- Estimated Net†††††
Population Density††
Population Diversity – Ethnic‡‡‡‡‡
Population Diversity – Religious‡‡‡‡‡
Population Growth††
Slum Population -- proportion of population†††††
Urban Growth Rate -- Annual percentage††
Youth Bulge -- Pop. Aged 0-14 as a % of Total††
6. Environment
Arable/fertile land availability††
Consumption -- Commercial energy consumption per capita††††
Consumption -- Use of solid fuels†††††
Disaster Risk Index§§§§§
Ecological Footprint -- Global hectares per capita******
Water -- annual withdrawal§§§§
Water -- Renewable available per capita§§§§
Forest -- annual percentage change in area§§§§
Pollution -- CO2 Emissions per capita††
Pollution -- CO2 Emissions per dollar PPP††
Energy--use of combustibles††

Table 20: Cluster List - Demography and Environment

Table 21: Supplemental Data for InterCoder Reliability

Event	CA+CE+ES	Sign	Causality	Centrality	Escalation	Cluster	ALC
1	-3	-1	3	3	2	1	1
1	-9	-1	3	3	3	3	1
1	-2	-1	2	2	2	3	1
1	-3	-1	3	3	3	3	1
1	-8	-1	2	3	3	3	1
1	-3	-1	3	2	2	1	3
2	-8	-1	3	2	3	3	1
2	-6	-1	2	2	2	3	1
2	-6	-1	2	2	2	3	1
2	-6	-1	2	2	2	3	2
2	-6	-1	2	2	2	3	3
2	-7	-1	3	2	2	3	3
3	7	1	2	2	3	3	1
3	3	1	1	1	1	3	1
3	7	1	2	2	3	3	1
3	6	1	1	2	3	3	1
3	6	1	2	2	2	1	3
3	5	1	2	2	1	1	3
4	-8	-1	3	2	3	1	1
4	-8	-1	3	2	3	1	1
4	-7	-1	2	2	3	2	1
4	-8	-1	2	3	3	1	1
4	7	1	2	3	2	3	1
4	8	1	3	3	2	1	3
5	-5	-1	2	1	2	1	1
5	-5	-1	2	1	2	1	1
5	-4	-1	1	1	2	1	1
5	-6	-1	2	2	2	1	3
5	-4	-1	1	2	1	1	3
5	-5	-1	1	1	1	1	3
6	4	1	1	1	2	1	1
6	4	1	1	1	2	1	1
6	5	1	2	2	1	1	2
6	3	1	1	1	1	2	2
6	4	1	1	2	1	2	2
6	5	1	2	2	1	2	2
7	-4	-1	1	2	1	3	1

7	-6	-1	2	2	2	3	1
7	-5	-1	1	2	2	3	1
7	-6	-1	2	2	2	1	3
7	-6	-1	2	2	2	1	3
7	-6	-1	2	2	2	1	3
8	6	1	2	2	2	1	3
8	7	1	2	3	2	1	3
8	6	1	2	2	2	1	3
8	7	1	2	3	2	1	3
8	6	1	2	2	2	1	3
8	5	1	1	2	2	1	3
9	-6	-1	2	2	2	2	1
9	-8	-1	3	3	2	3	1
9	-8	-1	3	2	3	3	1
9	-7	-1	2	2	3	3	1
9	-8	-1	3	2	3	3	1
9	-8	-1	3	3	2	3	2
10	-3	-1	1	1	1	1	1
10	-3	-1	1	1	1	1	1
10	-6	-1	2	2	2	4	2
10	-3	-1	1	1	1	4	2
10	-5	-1	1	1	3	4	2
10	7	1	3	2	2	5	2
11	4	1	2	1	1	2	1
11	7	1	2	3	2	1	1
11	5	1	2	1	2	3	1
11	5	1	2	2	1	3	1
11	4	1	2	1	1	3	1
11	6	1	2	2	2	1	3
12	7	1	2	3	2	1	1
12	7	1	2	3	2	3	1
12	7	1	2	2	3	3	1
12	6	1	2	2	2	3	1
12	5	1	2	2	1	3	1
12	-5	-1	2	2	1	5	3
13	-5	-1	2	2	1	1	1
13	-3	-1	1	1	1	1	1
13	-5	-1	2	1	2	1	1
13	-4	-1	1	2	1	1	3
13	-4	-1	1	1	2	1	3
13	-6	-1	2	2	2	1	3

Table 22: Group 1 T-tests

T-tests : STABLE (Group 1) vs. UNSTABLE (GROUP 2) Years significant at p < .05											
	Mean 1	**Mean 2**	**t-value**	**df**	**p**	**N 1**	**N 2**	**SD 1**	**SD 2**	**F-ratio**	**p**
O_NUM	21.31	22.49	-2.97	1198	0.00	600	600	7.15	6.68	1.15	0.09
O_AVG	0.36	-0.20	5.83	1198	0.00	600	600	1.57	1.75	1.24	0.01
O_SD	4.03	4.22	-3.52	1198	0.00	600	600	0.88	0.97	1.21	0.02
O_-9	0.06	0.12	-2.25	1198	0.02	600	600	0.34	0.54	2.58	0.00
O_%-9	0.33	0.46	-1.22	1198	0.22	600	600	1.75	2.00	1.30	0.00
O_%<0	43.76	48.73	-4.74	1198	0.00	600	600	17.73	18.58	1.10	0.25
O_%<-5	8.07	14.53	-7.71	1198	0.00	600	600	11.12	17.26	2.41	0.00
A_% all	45.16	43.06	1.59	1198	0.11	600	600	23.18	22.65	1.05	0.56
A_Avg	0.31	0.02	1.93	1198	0.05	600	600	2.46	2.67	1.19	0.04
A_SD	3.67	3.67	-0.01	1198	1.00	600	600	1.70	1.87	1.22	0.02
A_-9	0.05	0.08	-1.64	1198	0.10	600	600	0.30	0.40	1.77	0.00
A_%-9	0.25	0.31	-0.74	1198	0.46	600	600	1.55	1.52	1.04	0.65
A_%<0	21.08	21.08	-0.00	1198	1.00	600	600	17.84	17.96	1.01	0.86
A_%<-5	5.42	8.35	-4.77	1198	0.00	600	600	8.73	12.30	1.98	0.00
L_% all	25.42	28.77	-3.06	1198	0.00	600	600	18.64	19.26	1.07	0.42
L_Avg	0.55	-0.31	5.73	1198	0.00	600	600	2.57	2.59	1.02	0.78
L_SD	2.69	3.27	-4.95	1198	0.00	600	600	2.07	1.99	1.08	0.37
L_-9	0.01	0.03	-2.61	1198	0.01	600	600	0.07	0.26	13.17	0.00
L_%-9	0.02	0.12	-2.66	1198	0.01	600	600	0.28	0.89	9.96	0.00
L_%<0	10.36	14.95	-6.55	1198	0.00	600	600	10.70	13.42	1.58	0.00
L_%<-5	1.66	4.20	-7.19	1198	0.00	600	600	4.08	7.67	3.54	0.00
C_% all	29.41	28.17	1.05	1198	0.30	600	600	20.95	20.12	1.08	0.32
C_Avg	0.50	0.12	2.66	1198	0.01	600	600	2.33	2.66	1.30	0.00
C_SD	2.89	2.93	-0.36	1198	0.72	600	600	1.90	2.04	1.14	0.10
C_-9	0.01	0.01	0.58	1198	0.56	600	600	0.12	0.08	2.00	0.00
C_%-9	0.06	0.02	1.05	1198	0.29	600	600	0.73	0.30	5.90	0.00
C_%<0	12.32	12.69	-0.54	1198	0.59	600	600	11.79	11.91	1.02	0.80
C_%<-5	0.99	1.97	-4.72	1198	0.00	600	600	3.16	4.00	1.60	0.00
G_% all	40.23	37.68	2.08	1198	0.04	600	600	21.52	20.87	1.06	0.45
G_Avg	1.10	0.44	5.12	1198	0.00	600	600	2.07	2.41	1.35	0.00
G_SD	3.30	3.50	-2.01	1198	0.04	600	600	1.70	1.77	1.09	0.30

T-tests : STABLE (Group 1) vs. UNSTABLE (GROUP 2) Years significant at p < .05											
	Mean 1	**Mean 2**	**t-value**	**df**	**p**	**N 1**	**N 2**	**SD 1**	**SD 2**	**F-ratio**	**p**
G_-9	0.01	0.05	-2.17	1198	0.03	600	600	0.15	0.39	6.94	0.00
G_%-9	0.05	0.18	-2.15	1198	0.03	600	600	0.56	1.40	6.30	0.00
G_%<0	14.90	16.67	-2.17	1198	0.03	600	600	13.53	14.61	1.17	0.06
G_%<-5	2.38	4.70	-5.50	1198	0.00	600	600	5.53	8.67	2.46	0.00
M_% all	17.52	16.01	1.82	1198	0.07	600	600	15.30	13.45	1.30	0.00
M_Avg	1.25	1.14	0.71	1198	0.48	600	600	2.43	2.76	1.29	0.00
M_SD	1.95	1.97	-0.13	1198	0.89	600	600	2.04	2.18	1.14	0.10
M_-9	0.00	0.00	0.00	1198	1.00	600	600	0.04	0.04	1.00	1.00
M_%-9	0.01	0.01	0.03	1198	0.98	600	600	0.16	0.16	1.08	0.34
M_%<0	4.77	4.27	1.40	1198	0.16	600	600	6.54	5.85	1.25	0.01
M_%<-5	0.56	0.90	-2.22	1198	0.03	600	600	2.50	2.75	1.21	0.02
S_% all	32.94	36.06	-2.27	1198	0.02	600	600	23.77	23.87	1.01	0.92
S_Avg	-0.98	-1.29	2.17	1198	0.03	600	600	2.59	2.47	1.10	0.26
S_SD	3.14	3.58	-3.81	1198	0.00	600	600	2.02	1.98	1.04	0.61
S_-9	0.04	0.07	-1.32	1198	0.19	600	600	0.27	0.34	1.50	0.00
S_%-9	0.25	0.26	-0.14	1198	0.89	600	600	1.58	1.29	1.50	0.00
S_%<0	19.61	22.92	-3.46	1198	0.00	600	600	16.10	17.03	1.12	0.17
S_%<-5	4.67	7.98	-5.65	1198	0.00	600	600	7.71	12.07	2.45	0.00
H_% all	7.15	7.99	-1.59	1198	0.11	600	600	8.48	9.74	1.32	0.00
H_Avg	0.23	0.01	1.42	1198	0.16	600	600	2.63	2.82	1.14	0.10
H_SD	1.14	1.50	-2.93	1198	0.00	600	600	1.93	2.25	1.36	0.00
H_-9	0.01	0.00	0.45	1198	0.65	600	600	0.07	0.06	1.50	0.00
H_%-9	0.02	0.01	0.66	1198	0.51	600	600	0.30	0.20	2.29	0.00
H_%<0	3.14	3.49	-1.18	1198	0.24	600	600	5.29	5.13	1.06	0.46
H_%<-5	0.31	0.70	-3.29	1198	0.00	600	600	1.87	2.20	1.38	0.00
D_% all	0.56	0.65	-0.74	1198	0.46	600	600	2.26	1.97	1.31	0.00
D_Avg	-0.04	-0.20	1.99	1198	0.05	600	600	1.18	1.57	1.78	0.00
D_SD	0.04	0.07	-1.06	1198	0.29	600	600	0.43	0.59	1.87	0.00
D_-9	(no events)										
D_%-9	(no events)										
D_%<0	0.30	0.43	-1.41	1198	0.16	600	600	1.64	1.56	1.11	0.19
D_%<-5	0.05	0.11	-1.59	1198	0.11	600	600	0.50	0.75	2.27	0.00

T-tests : STABLE (Group 1) vs. UNSTABLE (GROUP 2) Years significant at p < .05											
	Mean 1	**Mean 2**	**t-value**	**df**	**p**	**N 1**	**N 2**	**SD 1**	**SD 2**	**F-ratio**	**p**
E_% all	1.61	1.61	-0.02	1198	0.98	600	600	3.15	3.60	1.31	0.00
E_Avg	-0.32	-0.25	-0.65	1198	0.52	600	600	1.90	1.84	1.07	0.41
E_SD	0.23	0.19	0.70	1198	0.49	600	600	0.98	0.90	1.20	0.03
E_-9	0.00	0.00	0.00	1198	1.00	600	600	0.04	0.04	1.00	1.00
E_%-9	0.01	0.01	0.17	1198	0.86	600	600	0.16	0.13	1.64	0.00
E_%<0	1.04	0.95	0.63	1198	0.53	600	600	2.38	2.61	1.21	0.02
E_%<-5	0.08	0.14	-1.45	1198	0.15	600	600	0.56	0.91	2.60	0.00

Table 23: Correlation of CIFP and VRA Data

Correlations are significant at p < .05 N=768			
	1 – CS	**Hostile**	**Goldstein**
O_AVG	-0.19	-0.27	0.32
O_SD	0.26	0.29	-0.30
O_-9	0.12	0.11	-0.13
O_%-9	0.11	0.10	-0.12
O_%<0	0.19	0.27	-0.32
O_%<-5	0.14	0.22	-0.27
A_Avg	-0.24	-0.24	0.30
A_SD	0.25	0.25	-0.26
A_-9	0.11	0.09	-0.11
A_%-9	0.09	0.07	-0.09
A_%<0	0.36	0.39	-0.43
A_%<-5	0.21	0.28	-0.31
L_Avg	0.04	0.02	0.02
L_SD	-0.04	-0.02	-0.01
L_-9	0.08	0.08	-0.09
L_%-9	0.07	0.07	-0.08
L_%<0	-0.14	-0.11	0.07
L_%<-5	-0.03	-0.00	-0.03
C_Avg	-0.04	-0.10	0.11
C_SD	-0.15	-0.13	0.15
C_-9	0.03	0.03	-0.05
C_%-9	0.05	0.05	-0.06

Correlations are significant at p < .05 N=768			
	1 – CS	**Hostile**	**Goldstein**
C_%<0	-0.13	-0.10	0.13
C_%<-5	-0.03	0.03	-0.04
G_Avg	-0.00	-0.07	0.11
G_SD	-0.13	-0.08	0.07
G_-9	0.07	0.08	-0.09
G_%-9	0.06	0.07	-0.09
G_%<0	-0.11	-0.05	0.02
G_%<-5	-0.00	0.07	-0.08
M_Avg	-0.10	-0.13	0.16
M_SD	-0.11	-0.12	0.12
M_-9	0.05	0.04	-0.05
M_%-9	0.05	0.04	-0.05
M_%<0	-0.15	-0.12	0.11
M_%<-5	-0.03	0.03	-0.04
S_Avg	-0.14	-0.11	0.15
S_SD	0.14	0.16	-0.16
S_-9	0.11	0.09	-0.10
S_%-9	0.10	0.08	-0.09
S_%<0	0.37	0.39	-0.43
S_%<-5	0.20	0.25	-0.30
H_Avg	-0.08	-0.05	0.07
H_SD	-0.10	-0.08	0.12
H_-9	-0.00	0.01	-0.02
H_%-9	-0.01	0.01	-0.02
H_%<0	-0.08	-0.05	0.09
H_%<-5	-0.07	-0.03	0.04
D_Avg	0.00	-0.01	0.01
D_SD	0.05	0.01	-0.01
D_-9			
D_%-9			
D_%<0	0.05	0.01	-0.00
D_%<-5	0.02	0.01	-0.01
E_Avg	0.01	-0.03	0.05
E_SD	-0.07	-0.08	0.07
E_-9	0.02	-0.01	-0.01

Correlations are significant at p < .05 N=768			
	1 – CS	**Hostile**	**Goldstein**
E_%-9	0.02	-0.00	-0.01
E_%<0	-0.09	-0.09	0.08
E_%<-5	0.01	0.01	-0.01

Bibliography

Azar, Edward. (1994) "Event-Data and Software for Early Warning, Journal of Ethno Development." v.4, i.1. Conflict and Peace Data Bank (COPDAB) - Developed by Edward Azar, this data bank contains over 300,000 event records from 1948 to mid-1990s.

Bates, R., Epstein, D., Goldstone, J., Gurr, T. R., Harff, B., Kahl, C., Knight, K., Levy, M., Lustik, M., Marshall, M, Parris, T., Ulfelder, J. and Woodward. M. (2003) *Political Instability Task Force Report: Phase IV Findings*, McLean, VA: Science Applications International Corporation.

Bond, Doug, J. Craig Jenkins, Charles L. Taylor and Kurt Schock. (1997) "Mapping Mass Political Conflict and Civil Society: The Automated Development of Event Data." *Journal of Conflict Resolution* 41,4: 553-579.

Bond, Doug; Bond, Joe; Oh, Churl; Jenkins, J.C; Taylor, C. L. (2003) 'Integrated Data for Events Analysis (IDEA): An Event Typology for Automated Events Data Development', *Journal of Peace Research*, 40(6): 733-45.

Carment, D. (2003) "Assessing State Failure: Implications for Theory and Policy," *Third World Quarterly*, 24(3): 407-427.

Carment, D. and Garner, K (1998) "Conflict Prevention and Early Warning: Problems, Pitfalls and Avenues for Success," *Canadian Foreign Policy* (Winter): 103-118.

Carment, D. and Schnabel A. (eds) (2004) *Conflict Prevention, From Rhetoric to Reality: Opportunities and Innovations*, Lanham, MD: Lexington Press.

Carment, D., El-Achkar, S., Prest. S. and Samy, Y. (2006a) "The 2006 Country Indicators for Foreign Policy: Opportunities and Challenges for Canada," *Canadian Foreign Policy*, 13(1): 1-35.

Carment, D., Prest, S., Gazo, J., El-Achkar, S. and Samy, Y. (2006b) "Failed and Fragile States 2006: A Briefing Note for the Canadian Government," Country Indicators for Foreign Policy (June).

Carment, D., Prest, S., and Samy, Y. (2006c) "Assessing Small Island Developing State Fragility", in L. Briguglio, Cordina and E. J. Kisanga (eds.), *Building the Economic Resilience of Small States*, London: Formatek Publishing, on behalf of Islands and Small States Institute of the University of Malta and the Commonwealth Secretariat.

Carment, D., Gazo, J. and Prest, S. (2007) "Risk Assessment and State Failure," *Global Society*, 21(1): 47-69.

Carment, D., Samy, Y. and Prest, S. (2008a) "State Fragility and Implications for Aid Allocation: An Empirical Analysis," *Conflict Management and Peace Science*, 25: 349-373

Carment, D., Prest, S. and Samy, Y. (2008b) "Determinants of State Fragility and Implications for Aid Allocation: An Assessment Based on the Country Indicators for Foreign Policy Project," WIDER Research Paper No. 2008/46 (April).

Carment, D., Samy, Y. and Prest, S. (2008c) "Approaches to Country Risk Analysis and Early Warning," forthcoming in *Economia Internazionale.*

Carment, D., Prest, S. and Samy, Y (2009) *Security, Development and the Fragile State: Bridging the Gap Between Theory and Policy* (Routledge).

Clingendael, (2005) "The Stability Assessment Framework: Designing Integrated Responses for Security, Governance and Development", (January), available: <http://www.clingendael.nl/publications/2005/20050200_cru_paper_stability.pdf >.

CIFP Afghanistan Fragile States Brief , 2009 (www.carleton.ca/cifp).

CIFP Pakistan Governance and Democracy Report 2007 (www.carleton.ca/cifp).

CIFP Concept Paper 2005 (www.carleton.ca/cifp)

Davies, J. and McDaniel C. (1993) "The Global Event-Data System" in Merritt, R. L. Muncaster R. G., D and Zinnes, D.(eds.) *International Event-Data Developments: Ddir Phase II.*. Ann Arbor: University of Michigan Press.

Davies, J. (1994) "Event-Data and Software for Early Warning", *Journal of Ethno-Development*, 4,1.

DfID, "Why We Need to Work More Effectively in Fragile States", DfID (January 2005), available: <www.dfid.gov.uk/pubs/files/fragilestates-paper.pdf>.

Donner, Allan; Eliasziw, Michael. *Sample Size Requirements for Reliability Studies.* Statistics in Medicine. Vol 6, Issue 4. Pp441-448.

Esty, D., Goldstone, J., Gurr, G., Surko, P. and Unger, A. (1995) *Working Paper: State Failure Task Force Report*, McLean, VA: Science Applications International Corporation.

Esty, D., Goldstone, J., Gurr, T. R. Harff, B., Levy, M., Dabelko, G., Surko, P. and Unger, A. (1998) *State Failure Task Force Report: Phase II Findings*, McLean, VA: Science Applications International Corporation.

Events scaling. Online. Available HTTP: <http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/05211.xml > (accessed 10 September 2008).

FAST. Online. Available HTTP: <http://www.swisspeace.ch/typo3/en/peace-conflict-research/previous-projects/fast-international/about/index.html>. (accessed 1 July 2008).

Forum on Early Warning and Early Response (1998*) A Manual for Early Warning and Early Response* , London.

Fysh, A. (2007) *CIFP Fragile States Brief- Afghanistan*.. Online. Available HTTP:< http:www.carleton.ca/cifp> (accessed 1 January 2008)

Gerner, D. J. et al. (1994). "Machine coding of events data using regional and international sources," *International Studies Quarterly*, 38: 91-119.

Gerner, D. J. et al. (2002) “Conflict and Mediation Event Observations (CAMEO): A New Event Data Framework for the Analysis of Foreign Policy Interactions,” paper presented at the Annual International Studies Association.

German Federal Government, “Action Plan: ‘Civilian Crisis Prevention, Conflict Resolution and Post-Conflict Peace-Building’” (2004) 64, available: http://www.auswaertigesamt.de/www/en/infoservice/download/pdf/friedenspolitik/AP%20EN.pdf

Goldstein, J. (1992) “A conflict-cooperation scale for international events data,” *Journal of Conflict Resolution*. 36:369-385.

Goldstein, J., and Pevehouse, J. C. (1997) "Reciprocity, Bullying and International Cooperation: A Time-Series Analysis of the Bosnia Conflict." *American Political Science Review* 91,3: 515-530.

Goldstein, J. S., Pevehouse, J.C Gerner, D. J. and Telhami., S. (2001) "Dynamics of middle east conflict and U.S. influence, 1979-97," *Journal of Conflict Resolution*, 45,5: 594-620.

The Government of Canada. *Canada’s International Policy Statement: A Role of Pride and Influence in the World.* Her Majesty the Queen in Right of Canada: 2005.

Goldstone, J. et al.,(2000) *State Failure Task Force Report: Phase III Findings*, (McLean, VA: Science Applications International Corporation (SAIC), available: <http://www.cidcm.umd.edu/inscr/stfail/>.

Gros, J.G. (1996) “Towards a taxonomy of failed states in the New World Order: decaying Somalia, Liberia, Rwanda and Haiti”, *Third World Quarterly*, Vol. 17, No. 3. 1996. pp 455-471.

Gupta, Dipak. (1997) "An Early Warning About Forecasts: Oracle to Academics" in Susanne Schmeidl and Howard Adelman, (eds.), S*ynergy in Early Warning Conference Proceedings*, Toronto, Canada, March 15-18, 375-396.

Gurr, Tedd, R. and Monty G. Marshall, (2000) "Assessing Risks of Future Ethnic Wars." Chapter 7 and Appendix B in Ted Robert Gurr, *Peoples versus States: Minorities at Risk in the New Century.* Washington DC: United States Institute of Peace Press.

Gurr, Tedd and Harff, Barbara. “*Early Warning of Communal Conflicts and Genocide: Linking Empirical Research to International Responses”*. Tokyo: United Nations University. 1996

Hagmeyer-Gaverus, G. and Weismann, M (2003) Early Warning Indicators for Preventive Policy (SIPRI Working paper). Available at: www.carleton.ca/cifp.

Harff, B. (1998) “Early Warning of Humanitarian Crises: Sequential Models and the Role of Accelerators”. In L. Davies and T.R. Gurr. (eds.) *Preventive Measures: Building Risk Assessment and Crisis Early Warning Systems*, Lanham, MD: Rowman & Littlefield.

Hales, Doug; Miller, Jordan; Tikuisis, Peter. *Triggers of State Failure.* CAE Professional Services/DRDC Toronto-Adversarial Intent Section: 2008. CR 2008-054.

Huxtable, Phillip A. (1997) *Uncertainty and Foreign Policy-Making: Conflict and Cooperation in West Africa.* Ph.D. dissertation, University of Kansas.

Jenkins, J.C., Bond, D. (2001) “Conflict carrying capacity, political crisis, and reconstruction: A framework for the early warning of political system vulnerability,” *Journal of Conflict Resolution*. 45:3-31.

King, G. and Lowe, W. (2003) “An automated information extraction tool for international conflict data with performance as good as human coders: a rare events evaluation design.” *International Organization*, 57:617-642.

Klingebiel, S and Ogbamichael, H. (2004) “Poor Performers in Sub-Saharan Africa,” *African Security Review*, Vol. 13, no. 1 13-19.

Kovar, K., J. Fürnkranz, J. Petrak, B. Pfahringer, R. Trappl, and Widmer, G.. (2000) "Searching for Patterns in Political Event Sequences: Experiments with the KEDS Database," *Cybernetics and Systems* 31, 6.

Krummenacher, H. (2006) “Computer Assisted Early Warning - the FAST Example,“ in Robert Trappl (ed.) *Programming for Peace. Computer-Aided Methods for International Conflict Resolution and Prevention*. Dordrecht: Springer.

Marshall, M. and Goldstone, J. (2007) “Global Report on Conflict, Governance, and State Fragility 2007: Gauging System Performance and Fragility in the Globalization Era,” *Foreign Policy Bulletin* 17(1): 3-21.

McClelland, C. (1976) *World Event/Interaction Survey Codebook*. (ICPSR 5211). Ann Arbor, MI: ICPSR.

Merritt, Muncaster and Zinnes eds. (1993) "The Global Event-Data System," John Davies and Chad McDaniel in *International Event-Data Develop*ments.

Moreno Torres, M. and Michael Anderson, (2004) “Fragile States: Defining Difficult Environments for Poverty Reduction”, PRDE Working Paper, No. 1, (August 2004), available: <www.oecd.org/dataoecd/26/56 /21684456.pdf>.

Moore, W; Gurr, TR. *“Assessing risks of ethnopolitical rebellion in the year 2000: three empirical approaches’* in Schmeidl & Adelman. Synergy in Early Warning Conference Proceedings. September 1998.

Neumann, Lilly “Effects of scale length on means and correlation coefficients” *Quality and Quantity* (Vol. 17, no. 5, 1983).

O’Brien, Sean P. (2002) “Anticipating the Good, the Bad, and the Ugly: An Early Warning Approach to Conflict and Instability Analysis.” *Journal of Conflict Resolution* 46.6: 791-811.

O’Brien, Sean P. “Crisis Early Warning and Decision Support: Contemporary Approaches and Thoughts on Future Research” in *International Studies Review* (2010) 12, 87-104.

OECD, “Poor Performers: Basic Approaches for Supporting Development in Difficult Partnerships – Note by the Secretariat”, OECD DAC, (November 2001), available: <www.oecd.org/dataoecd/26/56/21684456.pdf>.

Pevehouse, J. C., and Goldstein, J. (1999) "Serbian Compliance or Defiance in Kosovo? Statistical Analysis and Real-Time Predictions," *Journal of Conflict Resolution* 43, 4: 538-546.

Prime Minister's Strategy Unit, "Investing in Prevention: An International Strategy to Manage Risks of Instability and Improve Crisis Response", A Prime Minister's Strategy Unit Report to The Government of the U.K. (February 2005), available: <www.strategy.gov.uk/work_areas/ countries_at_risk/index.asp>.

Raleigh, Clionadh; Ward, Michael D. *Prediction Models for African Conflicts*. 18 February 2010.

Rohloff, Kurt; Asal, Victor. *Computational Models to Discover Sets of Patterns of Behaviours that Precede Political Events.* Association for the Advancement of Artificial Intelligence: 2008.

Rotberg, R. (ed.) (2004) *When States Fail: Causes and Consequences*, (Princeton N.J.: Princeton University Press).

Schmeidl, S. and Bond, D.(2000) "Assessing the Implications of Automation in Event Data Analysis: An Examination of the Conceptual and Empirical Differences Between Human-Assisted and Fully-Automated Coding," paper presented at the Annual Convention of the International Studies Association, Los Angeles, USA, March..

Schrodt, P. A. (1999) "Early Warning of Conflict in Southern Lebanon using Hidden Markov Models," in Starr, H. (ed.) *The Understanding and Management of Global Violence*, New York: St. Martin's Press.

Schrodt, P. A. (2000) "Pattern Recognition of International Crises using Hidden Markov Models," in Richards, D. (ed.) *Political Complexity: Nonlinear Models of Politics,*. Ann Arbor: University of Michigan Press.

Schrodt, P. A., and Deborah J. Gerner. (2000). "Cluster-Based Early Warning Indicators for Political Change in the Contemporary Levant." *American Political Science Review* 94,4.

Schrodt, P. A. and Gerner, D. J. (1997) "Empirical Indicators of Crisis Phase in the Middle East, 1979-1995," *Journal of Conflict Resolution* 41,4: 529-552.

Schrodt, P. A. (2000). "Cluster-Based Early Warning Indicators for Political Change in the Contemporary Levant," *American Political Science Review* 94,4.

Schrodt, P. A., and Scott Savaiano. (1997) "Environmental Change and Conflict: Analyzing the Ethiopian Famine of 1984-85." pp. 147-158 in Carl W. Roberts, ed. *Text Analysis for the Social Sciences: Methods for Drawing Statistical Inferences from Texts and Transcripts.* Hillsdale, NJ: Lawrence Erlbaum Associates: 1997.

Sky, David, J. (2000) "Routine, Response, or Rational Expectations?: An Analysis of Bilateral Relations in International and International Contexts." Presented at the International Studies Association meetings, Los Angeles, 14-18 March 2000.

Suedfeld, Peter, Ryan W, Cross and Michael Stewart. Indicators, Predictors, and Determinants of Conflict Escalation and De-escalation: A Review of the Psychological Literature, Contract Report 2009072, Defence Research & Development Toronto, May 2009.

Trappl, Robert (ed). *Advances in Group Decision and Negotiation – Programming for Peace: Computer-Aided Methods for International Conflict Resolution and Prevention.* Dordecht, Netherlands: 2006

USAID (2005) Fragile States Documents. Washington DC.

US Department of Defense, Quadrennial Defense Review, February 2010, http://www.defense.gov/qdr/

Veen, Tim. *Event Data: A method for analyzing political behavior in the EU*. The Methods & Data Institute, Centre for International Crisis Management and Conflict Resolution: University of Nottingham. September 2008. Pp1-24

Walter, S.D.; Eliasziw, Michale; Donner, Allan. *Sample Size and Optimal Designs for Reliability Studies.* Statistics in Medicine. Vol 17, Issue 1. Pp101-110
Warren, Martin "The effects of scaling on the Correlation Coefficient A Test of Validity (193) Journal of marketing Research Vol. X.

World Events Interaction Survey; C. 'World Event/Interaction Survey (WEIS) Project, 1966–1978)', Third ICPSR Edition, Ann Arbor, MI: Inter-University Consortium for Political and Social Research).

Wyjad, K. 2007 CIFP Democracy and Governance Handbook (available at www.carleton.ca/cifp).

List of symbols/abbreviations/acronyms/initialisms

DND	Department of National Defence
ALC	Authority, Legitimacy and Capacity
CIFP	Country Indicators for Foreign Policy
PITF	Political Instability Task Force
VRA	Virtual Research Associates (VRA)
SME	Subject Matter Experts
EWM	Early Warning Model
FI	Fragility Index
PMESII	Political, Military, Economic, Social, Infrastructure, and Information
WEIS	World Events Interaction Survey

Distribution list

[i] VRA offre un outil informatique Web qui prend en charge l'analyse interactive et l'affichage intuitif de données de presse sous forme de tableaux, de graphiques et etc. Les affichages visuels représentent les avertissements de menaces pour les possibles « points chauds » internationaux.

DOCUMENT CONTROL DATA
(Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)

1. ORIGINATOR (The name and address of the organization preparing the document, Organizations for whom the document was prepared, e.g. Centre sponsoring a contractor's document, or tasking agency, are entered in section 8.)

Publishing: DRDC Toronto

Performing: CAE Professional Services 300–1135 Innovation Dr, Ottawa ON K2K 3G7

Monitoring:

Contracting: DRDC Toronto

2. SECURITY CLASSIFICATION
(Overall security classification of the document including special warning terms if applicable.)

UNCLASSIFIED

3. TITLE (The complete document title as indicated on the title page. Its classification is indicated by the appropriate abbreviation (S, C, R, or U) in parenthesis at the end of the title)

Indicators of State Failure: Phase II (U)
(U)

4. AUTHORS (First name, middle initial and last name. If military, show rank, e.g. Maj. John E. Doe.)

David Carment; Yiagadeesen Samy; Doug Hales; Jordan Miller; Peter Tikuisis

5. DATE OF PUBLICATION (Month and year of publication of document.)	6a NO. OF PAGES (Total containing information, including Annexes, Appendices, etc.)	6b. NO. OF REFS (Total cited in document.)
August 2010	103	80

7. DESCRIPTIVE NOTES (The category of the document, e.g. technical report, technical note or memorandum. If appropriate, enter the type of document, e.g. interim, progress, summary, annual or final. Give the inclusive dates when a specific reporting period is covered.)

Contract Report

8. SPONSORING ACTIVITY (The names of the department project office or laboratory sponsoring the research and development – include address.)

Sponsoring:

Tasking:

9a. PROJECT OR GRANT NO. (If appropriate, the applicable research and development project or grant under which the document was written. Please specify whether project or grant.)	9b. CONTRACT NO. (If appropriate, the applicable number under which the document was written.)
	CR 2010–124

10a. ORIGINATOR'S DOCUMENT NUMBER (The official document number by which the document is identified by the originating activity. This number must be unique to this document)	10b. OTHER DOCUMENT NO(s). (Any other numbers under which may be assigned this document either by the originator or by the sponsor.)
DRDC Toronto CR 2010–124	

11. DOCUMENT AVAILABILITY (Any limitations on the dissemination of the document, other than those imposed by security classification.)

Unlimited distribution

12. DOCUMENT ANNOUNCEMENT (Any limitation to the bibliographic announcement of this document. This will normally correspond to the Document Availability (11), However, when further distribution (beyond the audience specified in (11) is possible, a wider announcement audience may be selected.))

Unlimited announcement

DOCUMENT CONTROL DATA
(Security classification of the title, body of abstract and indexing annotation must be entered when the overall document is classified)

13. ABSTRACT (A brief and factual summary of the document. It may also appear elsewhere in the body of the document itself. It is highly desirable that the abstract of classified documents be unclassified. Each paragraph of the abstract shall begin with an indication of the security classification of the information in the paragraph (unless the document itself is unclassified) represented as (S), (C), (R), or (U). It is not necessary to include here abstracts in both official languages unless the text is bilingual.)

(U) As noted in the Canada First Defence Strategy and reiterated in the more recent US Quarterly Defence Review, instability and state failure in distant lands can directly affect our own security and that of our allies. Development of a predictive model has become both a topical issue and an increasingly important area of research in academic and policy communities. This is the second report documenting CAE's support to DRDC's continuing efforts to develop an Early Warning Model (EWM) of state instability.

The conceptual framework for an EWM was developed in a previous project, though without a data set to validate assumptions and the general hypothesis. The focus of the current project was to collect and code events data and integrate it with structural data that will ultimately be used to calibrate and validate the conceptual model. A descriptive framework was established and incidences of failure identified using the methodology developed by Carleton University's Country Indicators for Foreign Policy (CIFP) project. Events data were then collected for the 24 months preceding these instances of state failure. Periods of relative stability for these states were also identified and events data collected for these periods. These events were distinguished as eroding or bolstering state Authority, Legitimacy or Capacity, including the severity of the challenge recorded. An inter–coder reliability test was conducted to confirm coding consistency. The results were compared with data available through Virtual Research Associates (VRA) , thus affording an opportunity to gauge the merits of human (versus machine) coding. The research effort concluded that human coding is more discriminating but also considerably more time consuming. An extensive data base has been developed and analysis commenced, which will continue beyond the submission of this report.

The conceptual model envisages integrating events and structural data which would allow for the measurement and monitoring of state tension and, through regression analysis, for vulnerability and instability thresholds to be determined and crises of interest to be forecast. This report documents a uniquely extensive data base that has been developed to support this effort.

(U) Comme il est indiqué dans la Stratégie de défense Le Canada d'abord, et plus récemment encore dans le Quaterly Defence Review aux États–Unis, l'instabilité et la mise en déroute d'un État à l'autre bout du monde peuvent nuire directement à notre propre sécurité et à celle de nos alliés. Le développement d'un modèle prédictif est devenu un problème d'actualité et un sujet de recherche de plus en plus important pour le milieu de l'enseignement et pour le secteur des politiques. Le présent rapport est le deuxième à documenter le soutien de l'IAO aux efforts constants de RDDC pour le développement d'un modèle de pré–alerte (EWM) de l'instabilité des États.

Le cadre conceptuel d'un EWM a été développé au cours d'un projet antérieur, malgré l'absence d'un ensemble de données pour valider les présomptions et l'hypothèse générale. Le présent projet a été axé sur le recueil et le codage de données en vue de leur intégration à des données structurelles qui seront ultérieurement utilisées pour étalonner et valider le modèle conceptuel. Un cadre descriptif a été élaboré, et des occurrences d'échecs ont été cernées au moyen de la méthodologie développée par le Projet des indicateurs–pays pour la politique étrangère (CIFP) de l'Université Carleton. Des données d'événements ont été recueillies au cours des 24 mois qui ont précédé ces occurrences d'échecs d'États. Des données d'événements ont également été recueillies au cours des périodes de stabilité relative qui ont été cernées pour ces pays. Ces événements ont été classés selon qu'ils nuisaient ou contribuaient à l'autorité de l'État, à

sa légitimité et à sa capacité, y compris la gravité du défi enregistré. Un test de fiabilité d'inter−code a été fait pour s'assurer de l'uniformité du codage. Les résultats ont été comparés avec des données obtenues de Virtual Research Associates (VRA) , ce qui a permis de comparer le codage humain avec le codage machine. Les recherches ont mené à la conclusion que le codage humain est plus discriminatoire, mais aussi beaucoup plus laborieux. Une base de données exhaustive a été créée; son analyse a été commencée et elle se prolongera après la soumission de ce rapport.

Il est prévu que le modèle conceptuel servira à intégrer des événements et des données structurelles, ce qui permettrait de mesurer et de surveiller les tensions d'États. De plus, au moyen d'une analyse de régression, il serait possible de déterminer les seuils de vulnérabilité et d'instabilité, et de prévoir les crises d'intérêt. Une base de données exhaustive unique, qui a été développée en appui à ces recherches, est documentée dans le présent rapport.

14. KEYWORDS, DESCRIPTORS or IDENTIFIERS (Technically meaningful terms or short phrases that characterize a document and could be helpful in cataloguing the document. They should be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location may also be included. If possible keywords should be selected from a published thesaurus, e.g. Thesaurus of Engineering and Scientific Terms (TEST) and that thesaurus identified. If it is not possible to select indexing terms which are Unclassified, the classification of each should be indicated as with the title.)

(U) intra−state conflict;country instability; political violence; early warning; crisis forecast